Serial
TA350
.I42
v.7, 1956

CLEMSON COLLEGE LIBRARY

JUN 21 1957

CLEMSON, SOUTH CAROLINA

INDUSTRIAL mathematics

1956

Clemson College Library

INDUSTRIAL MATHEMATICS

THE JOURNAL OF THE
INDUSTRIAL MATHEMATICS SOCIETY

EDITED BY CHARLES LIPSON

VOLUME 7

1956

DETROIT WAYNE STATE UNIVERSITY PRESS

Copyright 1957 by the Industrial Mathematics Society,
100 Farnsworth, Detroit 2
Library of Congress Catalog Card Number 53-39635

EDITORIAL STAFF

Charles Lipson, University of Michigan, Editor-in-Chief
S. D. Conte, Wayne State University
R. V. Churchill, University of Michigan
K. W. Folley, Wayne State University
A. W. Jacobson, Wayne State University
R. N. Janeway, Chrysler Corporation
E. D. McCarthy, University of Detroit
C. A. Piper, Bendix Aviation Corporation
R. A. Roggenbuck, Ford Motor Company
S. Rosen, Wayne State University

SERIAL
TA350
.I42
V. 7, 1956

Composition and Lithoprinting by
BRAUN-BRUMFIELD, Inc.
Ann Arbor, Mich.

FOREWORD

The Industrial Mathematics Society is a professional organization whose objective it is to extend the understanding and application of mathematics in industry. Founded early in 1949, it marked the first organized attempt of any group to foster a closer relationship between mathematics and industry.

It promotes the cause of mathematics in industry through monthly meetings at which technical papers are presented, through lectures by nationally prominent scientists and engineers, through panel groups which concentrate in specialized fields of knowledge, and through publication of worthwhile papers in its annual volume.

Membership is open to engineers, scientists, and mathematicians who are concerned with the effective use of mathematics in industry. Annual dues are $5.00 per year including a subscription to the annual volume, *Industrial Mathematics*. Application forms and information may be obtained by writing to the Secretary, Industrial Mathematics Society, 100 Farnsworth, Detroit 2, Michigan.

146786

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

CONTENTS

		Page
1.	The Engineering Applications of Statistics . . . . . David N. Collins	1
2.	Determination of the Necessary Sample Size in Life-Testing When the Length of Life is Exponentially Distributed . . . . . Calvin J. Kirchen	17
3.	The Use of Computing Machines . . . . . Claus B. Ludwig	25
4.	Tire Vulcanization - An Application of Unsteady Heat Transfer . . . . . Samuel R. Schaffer	47
5.	Some Non-Parametric Methods of Testing the Validity of Assuming that a Sample Comes from a Given Population . . . . . Leonard G. Johnson	57
6.	Analytical and Experimental Studies of Vehicle Dynamic Behavior . . . . . John A. Carlson	77
7.	Improving the Validity of Transducer Measurements with Computing Circuits . . . . . S. A. Lippman	109
8.	Simplified Numerical Solution of a Catenary . . . . . Elihu Geer	119
9.	The Mathematical Background of Spiroid Gears . . . . . Oliver Saari	131
10.	Design of Orifices for Hydraulic Shock Absorbers . . . . . T. E. Lohr	145

1

The Engineering Applications of Statistics

By David N. Collins
Chrysler Corporation
Engineering Division
Detroit

The operation of any business, whether it is building automobiles or building a hamburger, brings an endless succession of questions. Can I use a different steel? Can it be lighter and still stand the strain? Did the butcher take me on that last batch of meat? The businessman's feeling of utter helplessness which accompanies any serious look at the day's list of questions has welcomed the help of engineering and research groups, or, lacking them, anyone at all who claims to have The Answer—any answer.

But engineers, to answer the questions of business, have questions of their own. How many service failures will occur with this new steel? If this part fails at 4,307 pounds, will the other 999,999 parts we plan to make also fail at 4,307 pounds, exactly? Are there any standards for hamburger? These questions get mathematical, and the branch of mathematics which can answer them best is statistics. It is the engineer's feeling of helplessness which adds the mathematical statistician to the payroll.

The addition of a statistician to the engineering staff may multiply the engineer's problems, rather than divide them. "What is a regression? Why won't you plot a smooth curve to show our data? What if it does have nine inflection points?"

The statistician has his own problems. "Don't these people understand statistical significance? How can they expect conclusive results from only one test, or even two?"

To sum up, there is a gap between the engineer and the statistician which must be bridged before the statistician is of much value to engineering. This paper will attempt to define the gap, and to suggest several bridges through which useful results can be obtained.

ENGINEERING ENDS AND STATISTICAL MEANS

Engineers use mathematics. Statisticians use mathematics. But the mathematics of the engineer is applied mathematics, tied to observations of physical quantities, and it must be resolved into physical effects. Statisticians use artificial arrays of numbers which may (or may not) be associated with physical observations. Luckily

these arrays and other techniques can greatly expand the range and sensitivity of engineering experiments. But it does not follow that every statistical operation has a physical counterpart.

ENGINEERS must make accurate and reproducible measurements, characterize and compare test methods, forecast responses to service conditions.

STATISTICIANS can separate and measure variances of groups of observations, measure the probability of real differences (not due to chance), evaluate constants in an equation of correlation.

The engineering applications of statistics are the use of the second set of functions for the benefit of the first.

ENGINEERING RESULTS FROM STATISTICAL METHODS

Probably the best non-mathematical link between statistics and engineering is through a discussion of individual applications for each method in turn. One of the simplest is the separation of variances within groups of observations; this is called the analysis of variance.

THE ANALYSIS OF VARIANCE

This is probably the most powerful single tool of statistics as applied to engineering. To illustrate, if a number of observations (of anything), made under varying conditions, are arranged in some kind of rectangular array, the result would be the vertical bars of Figure 1. These observations have an overall average (the plane of Figure 2) but are also divided by the rows and columns of the array into smaller sub-groups, each with its own average (Figure 3).

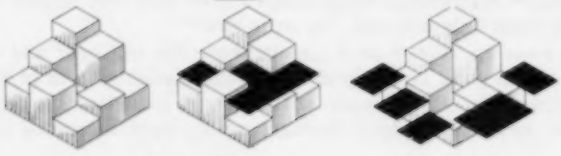


Figure 1 Figure 2 Figure 3

Each of these averages is accompanied by its calculated variance, which measures the variability of the observations represented by the average. So far, the arrangement of the array has been arbitrary, without relation to any physical quantity. But, if the rows and columns can be made to stand for physical conditions, then statistics and engineering come together with beneficial results that can be obtained in no other way. These quantities can be calculated:

1. The effect of each of the independent variables (usually one, two, or three) by which the observations were classified into rows, columns, and other groups.

2. The effect of each combination of variables taken two, three, or more, at a time.

3. The residual variation not accounted for by such a classification.

4. The probability that each of the effects found is due only to chance variation of the data. On this probability one can decide whether or not these effects are real and be right, say, nineteen times in twenty. In a word, the variance of all the observations has been analyzed into one section for each independent variable and combination of variables, with an unexplained remainder against which the others may be judged for chance differences. (Figure 4)

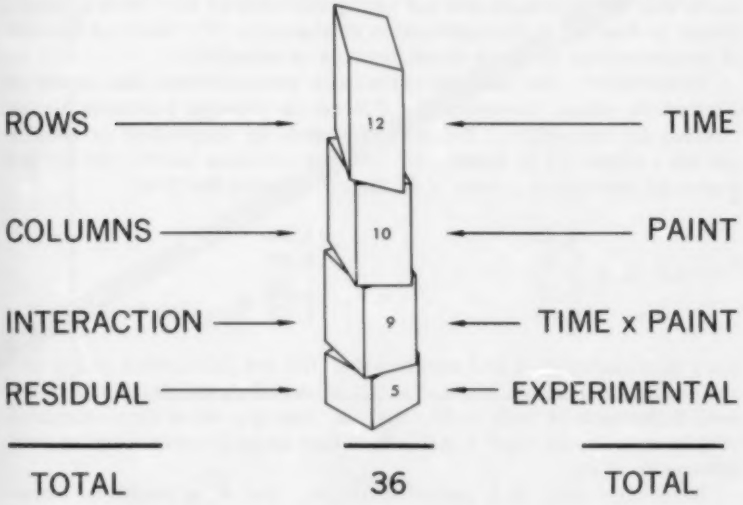


Figure 4

If the rectangular array is completely filled, that is, if there is one observation for each intersection of row and column, then this is a factorial design. There is an observation for each combination of the conditions assigned to the rows and columns. This is the statistical design which is most often compared with the classical design, which changes one variable at a time, while holding all the others constant. A detailed comparison of this type has been made by Duffett.[1] The advantage of the factorial over the classical design

is that less than twice as much work gives about four times as much information.

TESTING THE REALITY OF A DIFFERENCE

Probability calculations over the difference between two averages do not have the mathematical glamor of an analysis of variance (where "one problem makes a publication"), but they do provide a mathematical basis for every decision based on the reproducibility of results. The only numerical data required are the two averages and the variances associated with each average. If the difference between the averages is expressed in units of the combined variance, a figure can be found to express the probability that the difference is due only to chance variation, and that further measurement would show that the averages are not really different at all. Such a calculation is known as the application of student's "t", which is the unit of measurement in which the difference is calculated.

Practically, the use of calculated probabilities like these is limited to cases in which the difference between averages is not obvious by inspection. But the difference by inspection is usually not as obvious as it looks. Of the two columns below, the second shows an average increase of 61.9 percent over the first.

	4.48	11.01	
	3.59	5.27	
	6.42	7.16	
Av.	4.83	7.82	Increase 61.9%

Very straightforward and respectable. But the calculation of a probability figure through student's "t" shows that the probability of a real difference is only 0.66, that is, that one out of three samples of this size would show a difference this large if there were no real difference at all.

Naturally, this is a loaded example. But it is loaded no worse than much laboratory data from which differences are inferred and broad conclusions drawn. Remember that this "proof" of a difference still has a 66 percent chance of being confirmed by more data, but there is also a 33 percent chance that no difference exists. How much money would you bet on it?

Two things give this data a false appearance of authority. First, the results are calculated to two decimals, in spite of a variance of two to four units in each group. Second, the percentage increase looks impressive, although three samples which disagree so widely are too few to give a percentage difference any validity. Probability

calculations distinguish between guesses and proof.

CORRELATION

Statistics applied to graphical data produces measures of correlation. From any group of x, y observations the line or curve of best fit can be calculated, as well as a number which is a measure of the closeness of fit (Figure 5). The line or curve is calculated by the

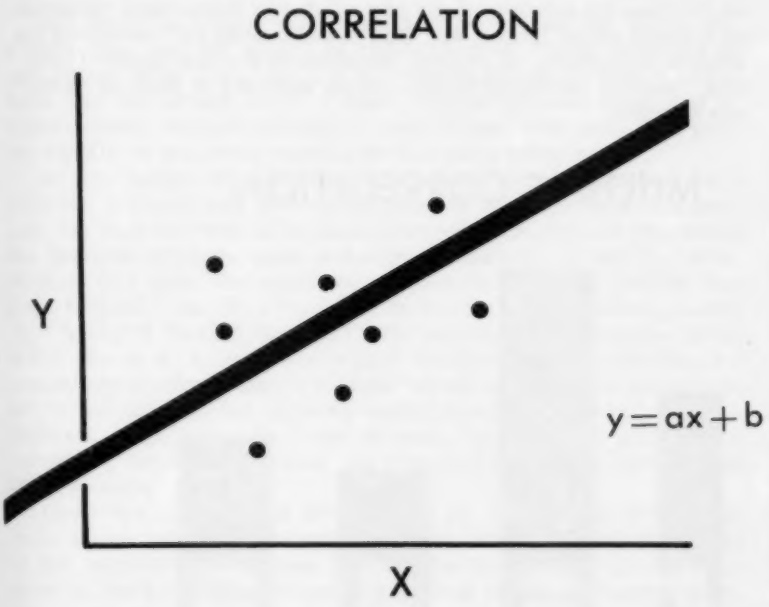


Figure 5

least squares method, which minimizes the sum of squares of the distances from each point to the calculated line. These are mathematical operations, of course, and the calculated line need not have any relation to a physical property of the data. A straight line, or a curve of any degree may be arbitrarily imposed on the data.

In the equation,

$$y = ax + b,$$

the calculated values are the slope, a, and the intercept, b. Although

the curves can be fitted mathematically, it is unusual to find anything but the straight line fitted in practice. It is easier to transform y into y^2, log y, or some similar function, and to fit the straight line than to go through the calculations for second degree or higher curves.

Another useful correlation which can be calculated is

$$y = ax_1 + bx_2 + cx_3 \ldots,$$

which gives the relative effects of a number of independent variables on a single result (Figure 6). This is most useful for measuring the relative importance of several variables on a single physical property, such as the effects of the alloying elements of steel on its tensile strength.

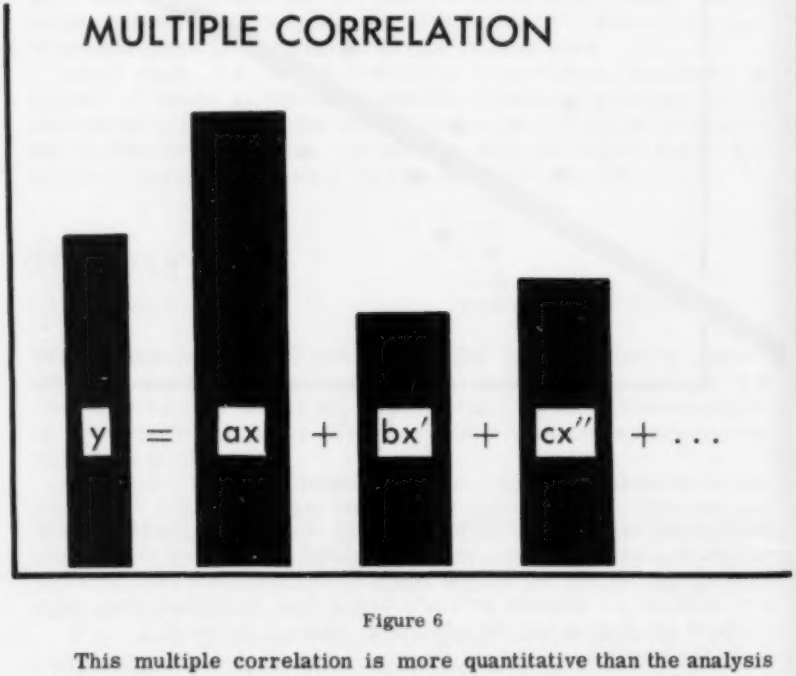


Figure 6

This multiple correlation is more quantitative than the analysis of variance which was mentioned earlier; where the usual analysis of variance only identifies the significant variables, multiple correlation evaluates their relative importance. Multiple correlation re-

quires a larger mass of data than the analysis of variance, and the calculations are more complex; however, the requirements for presetting the conditions of measurement are not as strict. The analysis of variance more easily shows interaction among the variables, which is only indicated in multiple correlation by a poor fit at some points between the observed data and the calculated line.

APPLICATION

DESIGN OF EXPERIMENTS

Not many technical men, these days, deny that the use of statistics improves the quality of research work. Most problems of application arise from the fact that the statistician is only called in after the laboratory work is done. Without pre-set conditions of measurement, statistics cannot be used to best advantage. The preset conditions are never reached without prior planning.

In the design of experiments, the statistician acts best as a referee, advising and arbitrating between two more important people—the man who will do the laboratory work and the man who asked the question which the work is designed to answer. In every conference of this type, the usual wall plaque which reads "THINK" (or even "THIMK") should be replaced by a more important one (Figure 7), "WHAT'S THE QUESTION?" The statistician's province is to tailor one of a stock of ready-made designs to fit the question, the laboratory worker makes his observations to carry out the design; but to set the question in terms which have laboratory answers is a problem which may take hours of work. Bicking[3] has published a check list for test programs; the required laboratory work is step 30 of 47 steps.

Once the question has been agreed on, two quantitative values must be set before the statistician's work begins. The first of these is the reliability of the test method; the second is the least difference in results which will be of practical value to the user. Also, this is the time to consider all the variables which might affect the results. Touchin's list[2] of the variables which affect exposure tests of paint panels is a good example; all these variables must be measured or controlled.

1. Preparation of the finishes.
2. The panels.
3. The human element in panel preparation.
4. The conditions during painting.
5. The conditions during drying.
6. The aging period before exposure.
7. Location of the exposure site.

8. Type of rack.
9. Position of the panel on the rack.
10. Weather immediately after exposure.
11. Weather over the complete period of the test.
12. Human element in evaluating breakdown.

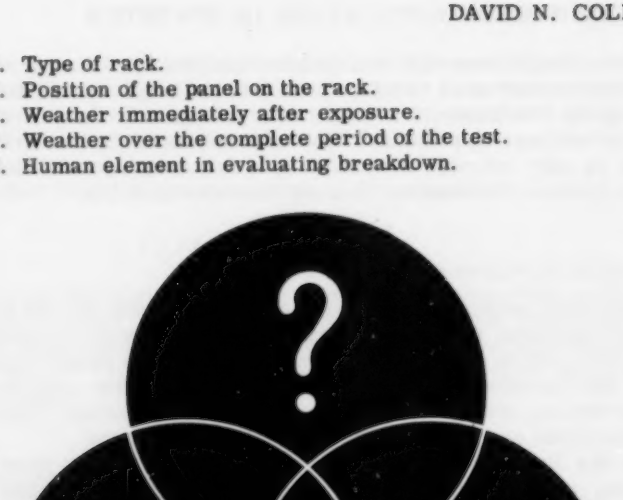


Figure 7

Without good planning, it is quite possible that valuable differences among paints (in this case) will be completely obscured by weather changes, different observers, and the like. An expensive program will have been spoiled for lack of good planning.

The outcome of a design conference should be the selection of a type of test program and the preparation of a detailed schedule of testing. The test schedule must be closely checked by the engineer, since he must attain in practice the experimental conditions which the statistician specified in theory. This does not mean that the

engineer should immediately reject as impractical all the unusual operating conditions which the test schedule demands, but a compromise must be reached between the impossible and the impotent.

Once the test schedule is set, it should not be changed in the slightest detail without the advice of the statistician. Small things are important—the interruption of work over a week-end or overnight; order differences in a routine operation—there will always be small points which were not covered in the plan. Most of these differences can be controlled and discounted if the statistician is aware of them beforehand, but some may be important enough to modify or destroy the whole program.

EVALUATION OF RESULTS

The biggest technical gap between the engineer and the statistician lies in the interpretation of results and the presentation of data. Statistical terms are mathematical, exact, and usually unintelligible; their engineering interpretations, to be equally exact, always include the element of chance that is expressed by a probability calculation. But the interpretations must be made, and the engineer must make them. Moreover, this area of uncertainty covers the one part of engineering research in which statistics gives the most help—sure-thing probabilities usually need no interpretation.

In a survey of statistical forms for the presentation of data, the most common one may be something of a surprise. It is a small thing, but its common adoption would double the amount of information available from laboratory work (Figure 8). It is this—that every average which represents a series of results should, as a matter of course, be accompanied by its standard deviation. Together, the average and the standard deviation are the best way to condense a series of measurements into two results. Alone, the average gives a false impression of accuracy; also, it is useless for evaluating the significance of differences.

Significant differences are expressed as probability fractions. A probability fraction of, say, 0.95 indicates that chance variation alone would account for the difference found, in a sample of the size taken, only five times per hundred samples, if the sampling is repeated indefinitely. Practically, this figure is used to imply that there is a 95 percent chance that the difference measured between two averages is due to a real effect of some changed condition, rather than to experimental error alone.

Several parts of the x, y chart of a series of measurements have statistical interpretations (Figure 9). The correlation coefficient, r, is a number expressing the degree of fit of the points to the cal-

Figure 8

culated line. When r = 1, the fit is perfect. If no line can be drawn through a cloud of completely random points, then r = 0. The r may be positive or negative without affecting its value, depending on whether the slope of the line is positive or negative.

Both the slope and the intercept of the x, y line are statistical quantities. Each has its calculated standard deviation which can be used to make comparisons with other slopes and intercepts. For example, Youden[4] uses an x, y line to calculate results in chemical analysis. Here the slope (grams CaO per gram of sample) is the

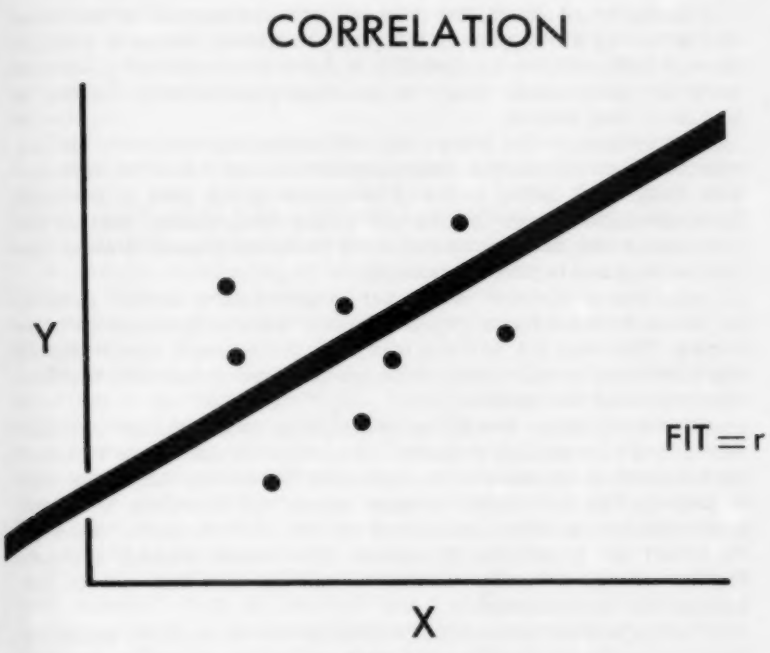


Figure 9

desired result, and the intercept is the "blank" error, if any. A probability figure for the existence of a blank can be obtained by calculating the probability that the intercept is different from zero.

The result of the analysis of variance is a table of variances, in which the effects of the rows, columns, and other subdivisions, and their interactions, are listed and compared with the experimental error. The table is packed full of information, but requires explanation to a user who is not familiar with this method of presentation.

HOW TO DISBELIEVE A STATISTIC

Statistics in research have reached the same degree of acceptance in the 1950's as business statistics did about 1820. These produced a series of loud explosions from British politicians and businessmen, like Prime Minister Canning's. "Egad, I can prove anything by statistics except the truth!" But, like today's researchers, the businessmen were forced to use statistics because those who used them were much better businessmen than those who did not.

Turning to research, the misuses and perversions of business statistics may also appear in the interpretation of research results. Darrell Huff, in a book called *How to Lie with Statistics*,[5] gives five qualifying questions for any data; these apply to corrosion results as well as to lung cancer.

WHO SAYS SO? The laboratory with something to prove for the sake of a theory is dangerously liable to the suppression of unfavorable data, or at least, to the presentation of the data in the most favorable light. Watch for the use of the "O.K. Name," that is, the reference to the well-known author or well-known method which has been misapplied beyond its capacity.

HOW DOES HE KNOW? Watch for evidence of a biased sample, one which does not fairly represent both sides of the question to be solved. This may not be conscious bias; for example, the choice of one panel from a pile never falls on the first or the last, but concentrates near the middle.

WHAT'S MISSING? Usually, what's missing is the number of trials on which the conclusion is based, or a reported difference in averages without a measure of variance on which the difference can be judged. The use of percentages without the raw data is a good issue clouder, whether intentional or not. For example, there is the report that in one year 33 percent of the women students at Johns Hopkins married faculty members. And so they had—one of the three students registered.

DID HE CHANGE THE SUBJECT? This is the first of two questions which cover the engineer's problem in evaluating statistical results. Watch the validity of the step which moves from results to conclusions. For instance, clocks A and B keep perfect time. When A shows nine o'clock, B strikes. Did A cause B to strike? These are two actions which are parallel, but which have no relation to each other.

Finally, the second of the engineer's questions, DOES IT MAKE SENSE? Watch for extrapolations, either above or below the range of measurement. Watch for calculating errors. In expressing variance, the standard deviation associated with an average is calculated to include two-thirds of the individual measurements within $\pm$ 1 s.d. of the average. Smaller or larger values are usually incorrect.

SALVAGE OPERATIONS

Once the data is taken, statistical evaluation is of little help except to condense the raw figures into manageable form. In a large testing program there is usually enough variety in the conditions to

permit several statistical comparisons, and the benefit of statistical planning is in the amount of work saved. However, most tests of average size have a comparison missing, or do not reduce the experimental error enough to prove real differences among the comparisons.

There are several methods of making the best of imperfect data. The most precise (and the least common) is to decide on the acceptable answer beforehand and select results to fit. Don't laugh; examine your own conscience. Without lecturing on professional ethics, I say this method is a complete waste of time and money.

A second method might be expressed with the words "Have Faith." Without enough data to prove the result, the conclusion is accepted as though it had been proved. This is the most common of all evaluations.

The faith of the engineer in his accuracy is the biggest single deterrent to statistical methods. Methods based on faith save laboratory time and money; and if the results are easily foreseen there is no great danger of making an incorrect conclusion. Even if the results are somewhat confusing, the conclusions may be qualified by use of the verbs "indicate" or "appear to indicate." In any case, a little laboratory work is better than none and gives an air of respectability to an otherwise unfounded conclusion.

There have been statistical salvage operations which produced good results. Both Bainbridge[6] and Lord[7] found their salvage so successful that the new conclusions were worth publishing for their technical value, as well as for a good example of statistical salvage.

Not all salvage work will be concerned with mistaken or ill-defined conclusions. Unless they are closely checked, electrical calculators give wrong answers with just as much speed and finesse as they give the right ones. Mistakes are made in transcribing figures from calculation to calculation. When conclusions depend only on numerical results, as these do, a small error can be deadly. Fyall's study[8] of the determination of size on glass fabrics contains four calculating errors in his basic data, all of which should have been caught and corrected by cross checking before the analysis of variance was started.

A LABORATORY UNDER CONTROL

Engineers in general have their troubles, and the suggestion that they should be burdened with statistical techniques would not be made except for two real improvements which follow. First, statistics save research time through elimination of the inconclusive project. Everyone can cite at least one such project—months or

years of data, progress reports which conclude, "More data will be required before final conclusions can be made." With lessening interest, there is the final report which sweeps all the crumbs of information into a pile, saying, "The data appear to indicate...," The data indicate nothing much, really. Ten thousand dollars spent for no conclusions.

The second benefit from statistical methods is that variance measurements of standard laboratory techniques can be calculated and reported and can be used for planning later work. This means smaller test plans, or increased precision, or both, as each test draws on previous results to prove significant differences. In a laboratory which uses all its data, the variance is as important as the average with which it is associated, and the report of an unusually high or low variance should raise as many eyebrows as the report that three Super Gas Savers in series will save 125 percent of your gasoline.

WHERE TO GET HELP

Such a revolution in laboratory operation will not come overnight, and will not come at all without expert advice. The best expert advice for day-to-day operations can be applied through the use of a statistical advisor at the lowest level of research planning. This advisor should not be a mathematician, but should be an engineer with statistical training who can apply his full time to advising on plans and to interpreting data. This is the "plant engineer" concept applied to research work—the man responsible for patching leaks and solving problems to keep up production. One man like this, plus a calculator operator, can handle the problems of 50-75 research workers.

The more-senior research planners who cannot spare time for training in the techniques may have to accept the techniques on faith for the results they bring. The technical literature of the last ten years is sprinkled with expository articles on "the new statistical approach," but none of them (including this one) have space enough to do more than suggest that statistics is a good thing and to illustrate some of the techniques. Do-it-yourself from books is slow, there is still no royal road to learning. However, most of the books on statistical methods will repay even a few minutes taken to read the Preface and Chapter I. Beyond this, the best results are obtained by classroom training.

Some training in statistical methods should be part of all scientific education. The half-year mathematics course which is the introduction to statistics seems quite satisfactory for teaching basic

techniques, but any application to the problems of laboratory work would be out of the question. This application, for chemists, should come quite casually in the undergraduate courses of quantitative analysis and physical chemistry, both in laboratory and classroom work. Some attention to the problems of variance in laboratory experiments would also help make laboratory training less formalistic and more practical.

For the graduate training of those who plan to specialize in the engineering applications of statistics, mathematics courses are already in use, built around the statistics of business. However, there should also be a graduate course in research planning to fit the special nature of laboratory statistics.

As more engineers are formally trained in statistics, there will be less of a gap between the engineer and the statistician. However, the demands on statistical specialists will increase, as present plans become more complex and new ones are developed to better fit the requirements of engineering research. As trained men become available, those of us can be retired who (like myself) have little talent for mathematics, but who are trying to fill the gap until better men arrive.

REFERENCES

1. J. R. Duffett, *Industrial Quality Control*, 11, No. 3 (1954), 36.
2. H. R. Touchin, *Journal, Oil and Colour Chemists Association*, 37 (1954), 248.
3. C. A. Bicking, *Industrial Quality Control*, 10, No. 4 (1954), 20.
4. W. J. Youden, *Statistical Methods for Chemists*, New York: John Wiley and Sons, 1951, Chap. 5.
5. D. Huff, *How to Lie with Statistics*, New York: W. W. Norton and Co., 1954.
6. J. R. Bainbridge, *Journal, Oil and Colour Chemists Association*, 35 (1952), 459.
7. A. Lord, *Journal, Oil and Colour Chemists Association*, 35 (1952), 437.
8. A. A. Fyall, *Determination of Size on Glass Fabrics*, British Royal Aircraft Establishment Technical Note Chemistry 1212, December 1953.

2

Determination of the Necessary Sample Size in Life-Testing When the Length of Life is Exponentially Distributed

By Calvin J. Kirchen, Engineering Staff
Ford Motor Company, Dearborn

FOREWORD

The necessary information for the determination of sample size for experimental or inspection work is given. Its use is exemplified in sampling inspection as an orientation. The major part of the paper is devoted to its application in life-testing work when times to failure have an exponential distribution. Formulas for the calculation of various quantities needed in this work and tables of important quantities are given.

At the beginning, it is necessary to state the basic concepts involved in sample-size problems. They will be stated in terms of the sampling inspection operation that is almost universally used in the receiving departments of our many industrial plants.

Suppose a shipment of 20,000 bolts is to be inspected. Since only a sample is to be inspected, it is not only possible but quite probable that there will be defective bolts which will not be picked up in the sample. In the event the shipment is accepted by sampling inspection, these defectives will remain in the shipment, to be discarded perhaps on the assembly line. The point of this is: requiring a shipment of 20,000 bolts to be 0 per cent defective is actually unrealistic, because not even 100 per cent inspection could guarantee it. Hence, some small per cent defective should be defined as "acceptable" per cent defective; this per cent defective is usually referred to as a "quality level," in particular, as the "acceptable quality level": AQL. Obviously, there must also be a specification of an "unacceptable quality level": UQL, and, of course, the UQL in per cent is greater than the AQL in per cent. It does not follow incidentally, that UQL is the minimum feasible per cent defective greater than AQL.

A little reflection on the fact that the sample may contain more than the expected number of defectives and hence indicate that the shipment of bolts being sampled is of unacceptable quality (an example of "sampling error"), when it is in fact acceptable, will make reasonable the concept that in drawing conclusions from sample data there is the unavoidable risk of rejecting an actually acceptable

shipment. This risk is always denoted by the Greek letter alpha, α, and is specified at the discretion of those administering the inspection operation It is customarily given as a percentage and rarely exceeds 10 per cent. It is important to associate AQL and α closely; in fact, they are inseparable. Similarly, there is associated with the UQL the risk of accepting a shipment of unacceptable quality; it is denoted by beta, β, and is likewise specified by the inspecting organization, with a value usually not exceeding 10 per cent. It may or may not be equal to α. A widely adopted system fixes α at 5 per cent and β at 10 per cent, but their values are arbitrary. In statistical language α is usually termed a Type 1 error and β a Type 2 error. In inspection circles they are known respectively as producer's risk and consumer's risk.

The relations of these four concepts are clearly shown by a graph which gives the probability of acceptance by a sampling plan of shipments of various grades of per cent defective. Such a curve is called the operating characteristic of the inspection plan. The operating characteristic of any inspection or experimental procedure for specified values of AQL, α, UQL and β can and should be calculated and graphed.

If we set up the following specifications for a sampling inspection plan: AQL = 2 per cent, UQL = 9 per cent, $\alpha = .05$ and $\beta = .10$, then standard statistical methods or a reference to F. L. Grubbs, "On Designing Single Sampling Inspection Plans," *Annals of the Institute of Mathematical Statistics*, 20, No. 2, 242-256, will show that inspecting a sample of 70 pieces taken at random from a lot and accepting if no more than 3 defective pieces are found, will satisfy the four requirements above. The operating characteristic is shown in Figure 1.

The remainder of this paper will be devoted to the use of these concepts in life testing, with particular emphasis on the case of an exponential distribution for the times to failure of items tested:

$$y = \frac{1}{L} e^{-x/L}$$

where x is measurement of life on some scale (units of time, or number of test-cycles) and L is average life of the material or kind of item tested. The methods described are based upon the paper "Life Testing" by Professors Epstein and Sobel, Wayne State University, in the *Journal of the American Statistical Association*, 48, No. 263 (1953), 486-502.

The basic concepts in terms of life testing are:

1. The minimum life (number of cycles of testing or life on time scale) a part must last to be acceptable: L_A. Associated risk of rejection is α.

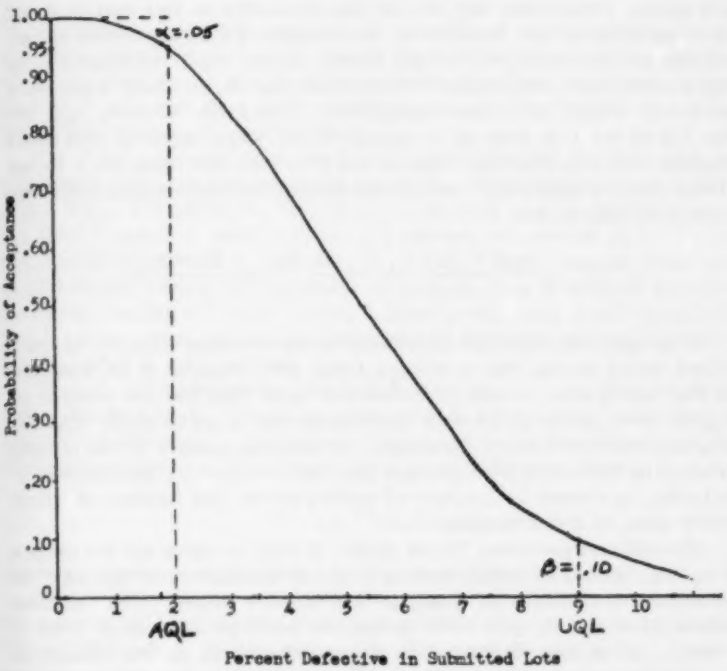


Figure 1

Operating Characteristic

2. The maximum life a part may last in being unacceptable: L_U. Associated risk of acceptance is β.

The notation given above (L_A and L_U) is different from that used by Epstein and Sobel, who used notation customary in mathematical statistics, which is frequently less suggestive of actual applications.

While it must be emphasized that an exponential distribution of times to failure is basic to the methods to be discussed, neither space nor time is available to discuss this further. See D. J. Davis, "An Analysis of Some Failure Data," *Journal of the American Statistical Association*, 47, No. 258 (1952), 113-150, with particularly the middle graph of the bottom three on p. 116 and the last paragraph on p. 117.

Professors Epstein and Sobel had particularly in mind the type of testing common with electron tubes where all items of the sample

are tested simultaneously so that the first tube to fail marks itself unmistakably as the weakest of the sample; the second tube to fail as the second weakest, and so forth. Under such conditions it is not necessary to test until all items have failed, to reach a decision as to acceptability or unacceptability. Using the notation $x_{i,n}$ for the life of the i-th item in a sample of n items, Epstein and Sobel showed that the statistic calculated from the life-data of r failed items out of n under test, which maximizes the exponential distribution given above, is:

$$L_m = \frac{x_{1,n} + x_{2,n} + \cdots + x_{r,n} + (n-r)x_{r,n}}{r}$$

It is apparent that this is similar to an average lifetime of the r failed items alone, but it differs from that because it is weighted by the final term. It can be remarked here that the calculation of L_m as given above is the *only* calculation that need be made with the $x_{i,n}$ in order to reach a decision. An unusual feature of the testing procedure that should be emphasized here is that r, the number of failures, is chosen in advance of testing and n, the number of items under test, is left arbitrary.

The natural question, "What guide, if any, is there for the choice of n, the number of items to be put under simultaneous testing?" is answered by Table 1 of Epstein and Sobel's paper. For various values of r and n, this table gives the average amount of time it takes to attain the rth failure in n compared with the rth failure in r. This table is calculated from:

$$\frac{\sum_{i=1}^{r} \frac{1}{n-j+1}}{\sum_{i=1}^{r} \frac{1}{j}}$$

An excerpt from Table 1 appears as follows:

r \ n	1	2	3	4	5	10	15
1	1	.50	.33	.25	.20	.10	.067
2	-	1.00	.56	.39	.30	.14	.092
3	-	-	1.00	.59	.43	.18	.12
4	-	-	-	1.00	.62	.23	.14
5	-	-	-	-	1.00	.28	.18

Assuming a rational basis for the choice of r, the experimenter has here a guide to making a choice from the many available values of n. Time-wise, it will be advantageous to make n somewhat larger than r; this will have to be balanced against the cost of obtaining items for test and for testing them.

The final phase of this work is the definition of the test of L_m whereby a decision can be made as to the acceptability of the items under test. Explicitly stated, what is desired is how "best" to use the first r ordered observations of a sample of n items in order to discriminate between L_A and L_U, $L_A > L_U$. "Best" has the meaning in statistical testing of hypotheses of specifying a test which has the property, among all tests having a fixed probability α of rejecting $L = L_A$ when true, of having the largest probability of rejecting $L = L_A$ when $L = L_U$ is true.

Methods of mathematical statistics are employed to obtain a test value denoted by C, which is expressed in terms of a well-known statistical distribution called chi-square, which is tabulated in most statistical texts, the acceptable life L_A, and the quantity r:

$$C = L_A \chi^2_{1-\alpha}(2r)/2r.$$

For acceptance of items under test, we must have $L_m > C$.

The factor which resembles "x square" is "chi-square." Both the subscript $(1-\alpha)$ and the 2r in parentheses require comment here. The shift from α to $1-\alpha$ was made for convenience in the mathematical statistical derivations. What is more important to the user is the fact that there is at least one set of tables of the chi-square distribution which defines the probability levels as complements of those Epstein and Sobel had in mind. This set is in A. Hald, *Statistical Tables and Formulas*, New York: John Wiley and Sons, Inc., which supplements their excellent work *Statistical Theory with Engineering Applications* (same publishers). After remarks on the meaning of (2r), a means of identifying the proper set of tables to use will be given. Chi-square tables are arranged in columns, one of which is headed by n (extreme left column) and the remainder by probability values. The entries in the n column are known as "degrees of freedom," and the quantity (2r) defines the number of degrees of freedom for which chi-square is evaluated. The number of components to be tested to destruction is, then, half the degrees of freedom. Since .05 is a frequent probability involved, it will be used to distinguish between Hald's table and those Epstein and Sobel had in mind; if the .05 probability column for 1 degree of freedom has a chi-square value of 3.84, then $\alpha = .05$ in Hald's sense. In Hald's chi-square tables, 3.84 appears in the .95 probability column for 1 degree of freedom.

For reasons of practicality both in theory and application, the ratio L_A/L_U or its reciprocal, and the value of L_A, are used rather than specific values of L_A and L_U. Values of r which are used are smallest integers such that

$$\chi^2_{1-\alpha}\,(2r)/\chi^2_{\beta}\,(2r) > L_U/L_A\,.$$

From the definition of the critical value of C above, we have

$$C/L_A = \chi^2_{1-\alpha}\,(2r)/2r\,.$$

Hence C can be evaluated in advance of experimentation. If L_m, calculated from the experimental data, exceeds C, then we conclude the material is acceptable, running the risk β of accepting material whose life is L_U. Epstein and Sobel made up nine sub-tables, using all combinations of .01, .05 and .10 for both α and β, and with L_A/L_U specified for the following series: 1.5, 2, 2.5, 3, 4, 5, and 10. In each sub-table can be read the appropriate value of r and C/L_A. Their table for $\alpha = .05$, $\beta = .05$ follows:

L_A/L_U	r	C/L_A
1.5	67	.8079
2	23	.6834
2.5	14	.6046
3	10	.5426
4	7	.4694
5	5	.3940
10	3	.2725

As an example of its use, if we assume $L_A = 100$ and $L_U = 20$, then $L_A/L_U = 5$ and hence $r = 5$. $C = L_A(.3940) = 100\ (.3940) = 39.40$. If we choose $n = 10$, then from Table 1 we see that 5 failures in 10 tested would be obtained on the average in .28 of the time required to observe 5 failures in 5 tested. This time ratio can be reduced even more if more than 10 are tested. At any rate, acceptance depends on $L_m > 39.40$. To illustrate the method of obtaining r and C for other than the given values of L_A/L_U, assume we want this ratio to be 8, with $\alpha = .05$, $\beta = .05$, so $L_U\ L_A = .125 \leqslant \chi^2_{1-\alpha}\,(2r)/\chi^2_{\beta}\,(2r)$. Comparison of the $1-\alpha = 1-.05 = .95$ column and the .05 column of a chi-square table shows that $n = 6$ is the least integer such that $.125 \leqslant 1.64/12.6$, so $2r = n = 6$, or $r = 3$, and $C/L_A = 1.64/6 = .27$.

As a final remark, the calculation of the operating charactertistic

of a life-testing experiment is shown by an example. The procedure is indicated by relation (27) in Epstein and Sobel's paper.

Probability of accepting $L = L_A$ when L is true value =

$$\text{Probability}\left[\chi^2(2r) > \frac{L_A}{L}\chi^2_{1-\alpha}(2r)\right].$$

Following the example of the operating characteristic shown for the sampling inspection plan, Figure 1, the probability of acceptance is plotted as the ordinate, with the ratio L_A/L as the abscissa. Let $\alpha = .05$, and a specific value of r be chosen; then $\chi^2_{1-\alpha}(2r)$ is a specific value in the chi-square table. Choose a value of L and obtain L_A/L. Multiply this by $\chi^2_{1-\alpha}(2r) = \chi^2_{.95}(2r)$. Find the value of $\chi^2(2r)$ equal to $\frac{L_A}{L}\chi^2_{.95}(2r)$. Note the probability level this value of χ^2 is in. Plot the probability level for the value of L_A/L. It will be found quickest to obtain a series of values for L_A/L by calculating $\chi^2(2r)/\chi^2_{.95}(2r)$ for the various values of $\chi^2(2r) > \chi^2_{.95}(2r)$.

3

The Use of Computing Machines

By Claus B. Ludwig
Motor Truck Engineering Department
International Harvester Company
Fort Wayne, Indiana

INTRODUCTION

One and a half centuries ago, the substitution for manual effort of heat energy for power plants inaugurated the steam age. Then nobody could foresee the incredible changes in business, production and machinery that would come into being as a result of the so-called industrial revolution. Today, industry is again experiencing fundamental changes which can be considered a new revolution. In order to meet present-day requirements of cost reduction and increased productivity, automatic control and data processing are now being employed in manufacturing and clerical operations. This trend toward increasing efficiency is growing with the improvement of machines which can perform the functions of control and data processing. An industry which does not recognize this trend will cease to exist.

Although self-control devices have been used for some time, only the discovery and fast development of electronics and its general application to feed-back systems have made it possible to improve control equipment enough to stimulate revolutionary changes in industry. The great evolution in the field of data processing is also due to electronics. During the war, new kinds of electronic computers were developed wherein the mechanical computing parts were replaced by magnetic and electrostatic units. The significance of these computers lay in their great flexibility and exceedingly high speed. It became possible to solve extensive mathematical problems in an unbelievably short time and to process tremendous amounts of statistical data. We know today that these electronic computers were used to solve such complex problems as those pertaining to the path of guided missiles and of projectiles for aviation defense, weather forecasting based on data from many stations, and atomic power development. The presence of complex problems in industry has also animated the application of the electronic computers in peacetime work.

This paper pertains to the application of automatic computers for design analysis and for the prevention of failures of machine parts.

It is, therefore, necessary to go into detail about automatic computers. In doing so, I have to talk about one of the most fascinating inventions made by man.

We all know of the two extremes public thought takes concerning these new electronic computers. One is that of disregarding most of the merits of the computers and reproaching the inventors for claiming their machines can "think." The other is that of waxing enthusiasm which claims that no problems will remain unsolved in the future and that "giant electronic brains" can do anything.

As always in life, the truth is somewhere in the middle. These machines have a great many merits, and, as we will see, their way of doing things can very well be regarded as thinking. On the other hand, these machines are far from being a "brain," although, the function of a brain is also thinking. To clarify, we can categorize the process of thinking into original thinking and routine thinking. Original thinking might make up, say, only 10 per cent of our mental effort and routine thinking the remaining 90. The general purpose of the electronic computers is to relieve men of the burden of the 90 per cent routine mental effort, but men must themselves still supply the important 10 per cent original mental effort to put the computers into use.

We engineers are now faced with a great variety of small and large-scale electronic computers and with the yet greater task of appraising them. We would like to know what these computers can do for us, what their limitations are, what effort we have to make in order to apply them most efficiently, and so forth. I want to discuss these points and to present an outline of engineering problems which may be solved by electronic computers.

PRESENT SITUATION IN ENGINEERING

The load and speed requirements of machine parts have increased tremendously in the past fifty years, while the size of machine parts has decreased. This development could not have resulted from sheer guesswork. It came about not only through practical experience but also through analytical work when it was impossible to find the optimum design by trial and error methods. Analytical work may be defined as the statement in mathematical terms of the behavior of material and related systems. The quality of this analytical approach depends upon the assumptions which are introduced into the theory. On the one hand, it is desirable to approach the exact theory as closely as possible by using accurate assumptions and exact boundary conditions. On the other hand, it is necessary to keep the assumptions and boundary conditions as

simple as possible in order to allow for a numerical solution of the problem. Therefore, the analytical approach is a compromise between the best theory available and the simplest theory possible.

In the first approach to a given problem, the simplified theory will outweigh the more exact theory and the answer may be equivalent only to a rough estimate. Then, as time and people improve the first approach, the answer becomes more and more refined until sooner or later the limit is passed beyond which the theory does not provide an economical answer, i.e., an answer which can be evaluated numerically in a reasonable time.

In the analytical approach, the language of mathematics is used to condense the entire concept of the investigation into a few short symbolic statements, but in the thorough investigation lengthy calculations may involve an enormous quantity of numerical data. I think it is safe to say that the mathematical solution of a given problem may require 10 per cent of mental effort in the sense of original thought and that the numerical evaluation of this solution may require the balance of 90 per cent of routine thought. Men will always be happy to spend the 10 per cent of original thought but the burden of 90 per cent routine thought kills nearly all of the more complex projects.

Now, in this situation, new helping tools are given to us which eliminate the monotonous and time-consuming computations and which can be the essential bridge between an engineering concept and its numerical solution. These new tools are the automatic computers which are completely controlled during operations of data processing by the use of stored programs in electrostatic memory or by the use of wired control panels.

SURVEY OF APPLICATION OF AUTOMATIC COMPUTERS

Electronic computers may be classified as either digital or analog computers. A digital computer is a device which operates with discrete values given in the form of digits. The accuracy of the performed result depends upon the number of digits the machine can handle. With the large scale digital computer, the precision can go up to approximately 150 digits. An analog computer can handle only physically measurable quantities. Therefore, the discrete values which are used in a digital computer have to be converted into quantities such as lengths, currents or angles for the purpose of calculating with an analog computer.

The difference between a digital and an analog computer can be explained by comparing a desk calculator and a slide rule, respectively. The slide rule is an analog computer because the logarithm

of numbers is represented by lengths. A desk calculator is a digital device because it can handle only discrete values in the form of digits. For both the analog and digital computer, the problem must be stated before it can be programmed. The difference is that the answer does not have to be known for an analog and the approximate answer must be known for a digital computer.

Analog Computer

In general, the analog computer can be applied in basic research studies. It will be used more for the investigation of the behavior of physical systems than for the evaluation of specific design data, since it has a limited accuracy. However, in problems where an accuracy with four digits is sufficient, an analog computer might be applied very efficiently. Of course, time will be required to translate the given information from the form of digital data into that of physical quantities. Also, the quality (i.e. the accuracy) of this translation depends upon the skill of the programmer. Figure 1 gives an outline of the applications of engineering problems occurring especially in the automotive industry soluble with an analog computer.

Figure 1

Analog Applications*

1. Parameter studies, that is, analysis based on insufficient data or where we must design by trial and error; the method involves the systematic substitution of different combinations of parameters which will be the most useful for achieving a desired result.
2. Vibration of structures. (For example torsional vibration in crankshaft.)
3. Spring systems and vibration absorbers.
4. Dynamics of mechanisms.
5. Fluid flow problems.
6. Analytical studies involving gears or other mechanical linkages requires the determination of backlash. Backlash angles from zero to any practical value are easily introduced to the system.
7. Engine crankshaft main bearing loads.
8. Lubrication of engines, where the problem is to determine the quantity and pressure of the lubricating oil at various points throughout the lubricating system.
9. Solution of sets of differential equations.
10. Gas flow in a fuel system and fuel combustion.
11. Behavior of valve in an engine.
12. Performance characteristics of engineering systems.
13. Simulation of thermodynamic cycles.
14. In the farm implement field, the design of machine parts, such as knife blades, to give the desired trajectory of the material being handled.

*Compiled by Miss H. S. Siegel of the International Harvester Company.

Digital Computer

In general, the digital computer can be applied for the numerical evaluation of mathematical formulae. Since the digital computer can execute only the four basic arithmetic operations of addition, subtraction, multiplication and division, all other operations, such as integrations, root extractions, differentations, evaluations of angles from trigonometric functions, have to be expressed in terms of these four basic operations. Therefore, considerable time is required to program given formulae. But we will see that, with a digital computer, the evaluation of a sine by using fifty terms of a power series, for instance, requires actually much less time than to look up this value in a table. Furthermore, the automatic computer eliminates the errors possible with the use of tables.

Generally speaking, there are two types of computations involved in technical problems:

1. Putting numbers into *formulae* which are either exact or close approximations.
2. Trial and error method when explicit formulae are replaced by a *set of formulae* and the calculations have to be repeated until the result is found.

Figure 2 gives an outline of the problems soluble with a digital computer by EDPM because of their monotonous part of numerical calculation. They are arranged into two groups as follows:

1. One group includes current problems which are handled at the present time by a desk calculator.
2. The other group consists of problems which were never attempted because of the lack of ability or time.

Obviously, this outline is meant only to be suggestive. Only experience will give a more complete list of all the possible engineering problems.

ELECTRONIC DATA PROCESSING MACHINES

From here on my paper will deal only with large-scale digital computers (EDPM) which can also be used as data processing machines. Examples of these automatic computers are the Univac by Remington Rand, CPC, 702 or 705 by IBM, Mark IV by Harvard University, Enivac and Edvac by Moore School of Engineering, and some others. These machines can be purchased or rented at different costs which may vary on the rental basis from $3,000 to $70,000 a month. Another possibility is to use a large-scale computer in one of the central computing laboratories at a rate of approximately $500 per hour.

Figure 2

Problems soluble with a digital computer by EDPM

Problems Solved by Desk Calculator	Problems Not Solvable at the Present Time
2 dimensional stress analysis	3 dimensional stress analysis
Cam design for single engine	General solution of cam problem
Torque converter for single engine	General solution of automatic transmission
Gear Calculations	
Vibration studies	
Spring designs	
Performance studies	
Bearing load calculations	
Crankshaft balancing	
	Market research studies
	Thermal conditions in an engine
	Distortions of valve seat, piston, head
	Muffler attenuation
	Manifold problem

The question arises: What type at which price shall anyone choose? This, of course, depends upon the size of the engineering outfit and the complexity of the problems which occur and also on whether or not the computer will be used for accounting purposes as material requirements and inventory control, maintenance problems, payroll jobs, etc. Careful consideration has to be given to all these factors before renting or purchasing one of these machines. General practice has been to send people to the schools conducted by the manufacturers of the computing machines to ascertain the feasibility of the machines for specific problems.

Short Description

As an example of a computing machine, the IBM 702 is shown in Figure 3.

Imput Units. Data are given into the machine from both magnetic tape and punch cards.

Output Units. Results can be written on magnetic tape, printed on paper, punched into cards, or plotted as curves from a television tube.

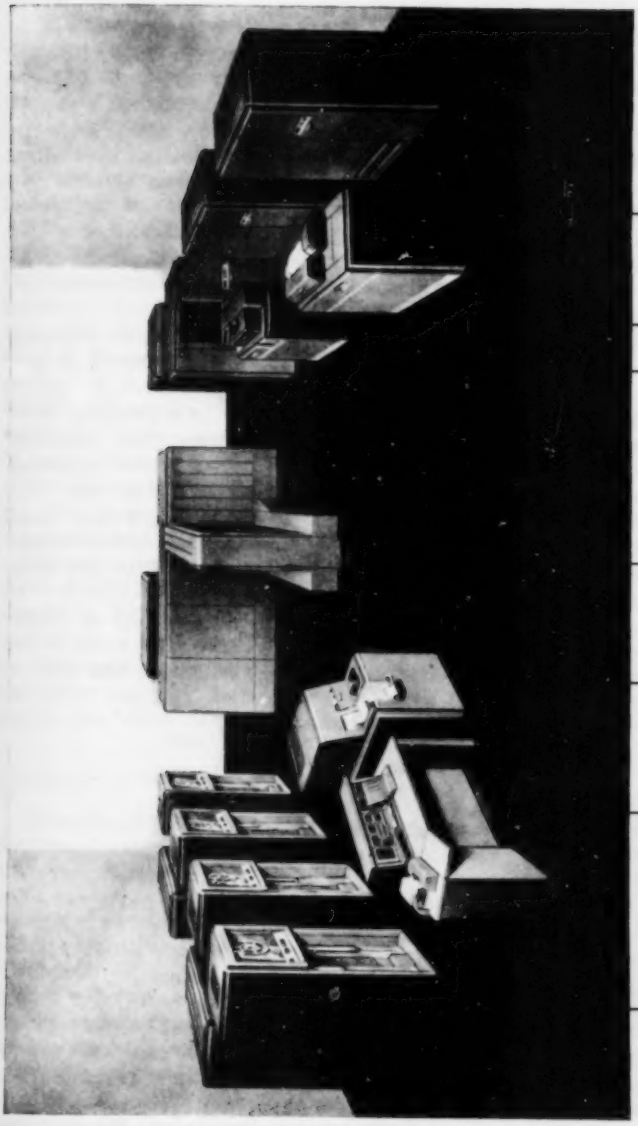


Figure 3

An Installation of IBM Electronic Data-Processing Machines Type 702 and Associated Equipment

Central Units. All mathematical operations are performed in the central unit which contains the electrostatic memory and the accumulators and the arithmetic and logical unit.

Drum Storage. Additional storage is possible with the use of magnetic drums.

Operations. Besides the four basic mathematical operations (addition, subtraction, multiplication, division), many logical operations can be executed by the machine. These include the modification, under given circumstances, of the instructions and decisions on such questions as the following: Are two figures equal, higher or lower? Is one figure minus, plus or zero?and Is the result right?

Internal Checking Procedures. Many possibilities of checking the internal operations are provided which give a control of accuracy of results.

Stored Program. The instructions included in a program contain five positions. The first position is the code of the operation which has to be performed. The remaining four positions represent the memory address at which the mathematical data is stored.

Since the program is stored in memory in the same way as data, the machine may modify instructions by substituting a new address. Therefore, a program can operate on itself and compute one of its own instructions.

Time of Operations. As an example of the high speed of the large-scale digital computers, the time required for some of the operations is of interest as it is given in the manual for IBM 702:

237,000	additions or subtractions (5 digits) per minute
50,000	multiplications (5 digits x 5 digits) per minute
24,000	divisions (5 digits) per minute
434,000	logical decisions per minute

The Limitations of the Large-scale Computer

The limitation of the digital computers is the human brain. The machine will never be more clever than the human brain which works with this machine. The machine will never be able to give a result which could not be reached by the human brain also. That is the limitation: What the human brain cannot figure out, the machine cannot either.

However, in many cases, the human brain would not be able to solve a problem in the time in which the answer is required. For example, the machine will not solve one equation with two unknowns because the human brain could not do it either. But the machine will calculate in one second the path of an airplane for air defense purposes where an individual would need several hours. Both answers will be right, but the human answer will be useless because of the time factor.

There is another interpretation of the statement, the limitation is the human brain. Not only is the machine no more clever than the human brain, but the human brain must be more clever than the machine. The human brain must have a definite concept of the problem which the machine shall solve. For example, one of the problems in our cam calculation is to find the most favorable power series. It is easy to make a program for the machine to calculate the 780,000 series. The machine would run about 1,000 hours and would give approximately 3 1/2 billion values. Without doubt, some of these values will be very favorable for our purposes, but who shall look for them? The great danger is that the machine could give too many values if there is no concept. These 1,000 hours machine time would be useless.

In order to make an intelligent use of the "electronic brain," there should be a definite concept of calculation for every single problem. This involves detailed theoretical investigations of the problems as well as programming. A program for one of these problems should never be based entirely on existing techniques but should be improved before programming for the EDPM. This will result in a better design and eliminate expensive test runs.

Preparation of Programs

In order to program a given problem for a digital computer, the problem must be stated in the language of the machine. This requires an exact knowledge of the whole problem. Even the approximate answer must be known. All equations and formulae have to be broken down into the four basic arithmetic operations. Then, in order to facilitate the programming, a detailed block diagram should be plotted, including all the steps of the machine operation. See Figure 4. The work up to this point has to be done by a skilled engineer who is well acquainted with the problem itself and with the digital computer as well. Anybody sufficiently familiar with programming can transfer the block diagram into actual program steps (i.e., the machine language). An average problem with approximately 1,000 instructions requires about two to four weeks for stating the problem and plotting the block diagram and about four days for writing the actual program.

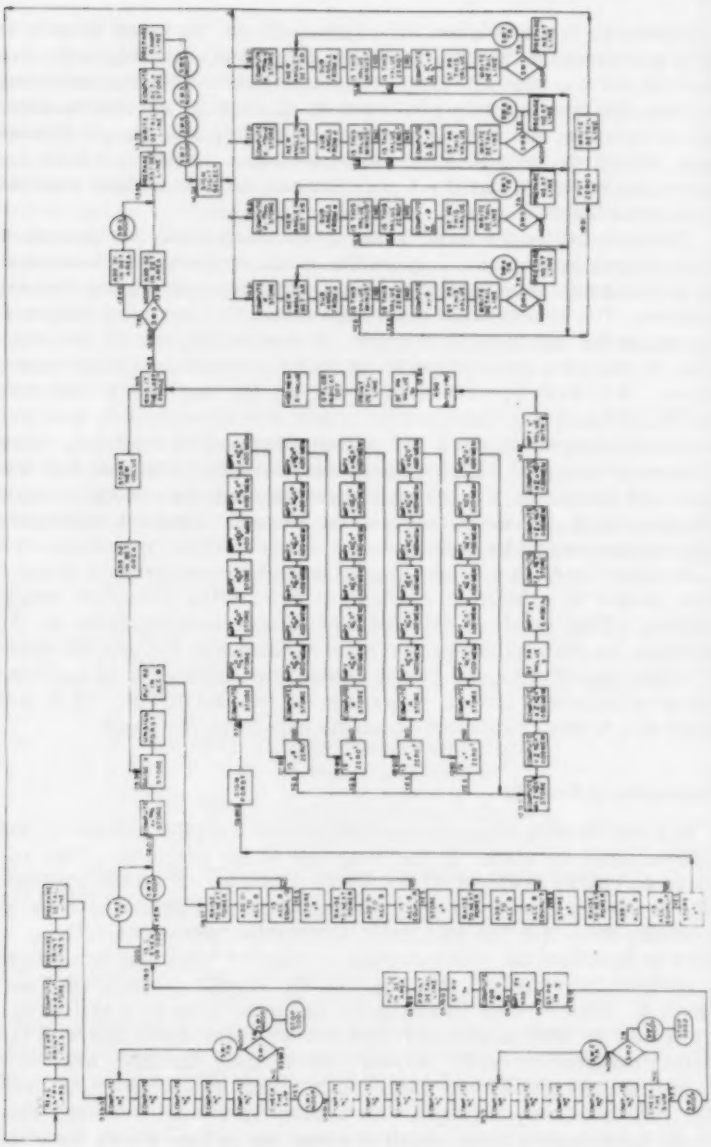


Figure 4

Detailed Block Diagram "Cam Calculation"
EDPM 702
CBL 003

Next, the written program is transferred onto punch cards (or whatever is needed for feeding into the machine). These punch cards can be stored to be used later when problems of the same nature arise, provided that the factors relating only to a specific design are put on separate cards (see Figure 5). Therefore, in writing the program, it is wise to plan for all future possibilities, i.e., to make the proper spacing. Nevertheless, the great flexibility even allows for future alteration in the basic program.

Another task is to prepare subroutines which include procedures and methods to be used over again in different problems. Examples of subroutines are extracting roots, obtaining the angles when the sines, cosines, etc., are given and vice versa, solving polynomials of any degree, etc.

The Staff as an Essential Supplement to the Machine

In order to fully utilize the large-scale digital computer for engineering projects, a definite program should be started in which a group of mathematician engineers become familiar enough with the potential of the equipment so that engineering problems may be applied effectively. This group has to accomplish two purposes:

1. To investigate all standard design procedures in which lengthy calculations are involved and to set up a new improved procedure. Many standard design problems can then be completely programmed and stored. If a new design is required, only one or two punch cards containing the specific design data and the code number of this program have to be transmitted to the machine. After a short time, the printed results will come back. The program remains in the storage room so that it is available for the next application. In this way, many design problems can be established within one hour while otherwise weeks of desk calculation would be required.
2. To investigate problems which cannot be solved economically at the present time. For these problems, in most cases much time is needed to find the most favorable approach.

Finally, it should be pointed out that the quality of the computer group determines the efficiency of the machine.

There appears to be a definite advantage in having a central group of engineers with the greatest possible competence instead of having several engineers from various sections for the following reasons:

1. The central handling of current design problems involving lengthy calculations will avoid overlapping.

EXPLANATION

1 MACHINE NUMBER
2 DATE
3 CAM NUMBER
4 SELECTED POWER SERIES
5 ROCKER ARM RATIO
6 HALF ANGLE OF EVENT
7 MAX VALVE LIFT
8 EQU. RAMP HEIGHT
9 EQU. RAMP VELOCITY
10 EQU. RAMP ACCELERATION
11 WEIGHT
12 SPRING RATE OF VALVE GEAR SYSTEM
13 SPRING RATE OF VALVE SPRING
14 CAM DESIGN SPRING
15 RAMP CALCULATION CODE

Figure 5
Detail Entry Card

2. Only a central group can coordinate the problems in such a manner that a great number of parts of EDPM programs can be reused for new problems.

3. The personnel in this group are kept in training for EDPM programming.

4. Because of their education, they will be able to increase the theoretical understanding of most of the given problems.

5. The people of this group will gain enough experience after a certain time so that they can apply the EDPM to unsolved problems most efficiently.

6. Because of their permanent training in analytical investigations, they can easily exploit their experiences in fields in which they were not previously familiar. They have the advantage of bringing a fresh approach to old problems; and, therefore, they might be able to find the general solution where only a small part was known before.

EXAMPLES OF APPLICATION

In this section, I will give a short survey of problems which have already been solved in the Motor Truck Engineering Department of the International Harvester Company.

First, it might be interesting to compare the actual machine time required by the different computers for one specific problem. Take, for example, the cam profile calculation based on the polydyne method. In this method, the evaluation of

$$y_o = h_r + \varrho_r(h + \sum_0 C_n\theta) + \phi \sum_n n(n-1)C_n\theta^{n-2}$$

has to be made where h_r, h, ϱ, ϕ and C_n are the design data. θ is the angle of duration and n indicates the powers p, q, r, s, t. There are approximately (all operations are executed with 7 digits)

3,000	additions and subtractions
4,000	multiplications
800	divisions and
1,000	logical decisions

involved. The machine time to execute these operations was on the

Desk Calculator	100 hours
IBM 604	12 hours
CPC	4 hours
IBM 702	20 seconds.

It appears that the time required for execution of this program for the different machines decreased tremendously as we used

more competent machines. This improvement is not only due to the higher internal speed but also to the ability of the larger machines to make logical decisions.

I do not have available the exact figures of rental cost, but I do know that the IBM 702 is much more expensive than the CPC. Therefore, a question arises as to whether or not the time saving of roughly four hours is worth the additional money for renting the IBM 702. The decision is in favor of the larger IBM 702 for the following reason. There are many programs which require more time than the twenty seconds on the IBM 702. If we assume a linear relationship between the machine times for the two computers, although actually the CPC will slow down for more complex programs, we have a time table. To illustrate, I will choose the evaluation of the entire solution of the differential equation which describes the valve gear train at non-design speed:

$$\frac{d^2y}{d\theta^2} + \beta^2 y = \beta^2\left[h + \sum_n C_n\theta^n\right] + \left(\frac{\beta}{\beta_0}\right)^2 \sum_n n(n-1)\,C_n\theta^{n-2}$$

This solution was obtained on the IBM 702 after about fifteen minutes. Assuming a linear relationship, the CPC would have needed 180 hours. It appears that for more complex problems, the smaller computers are no longer desirable for engineering applications.

Let me describe briefly the answer obtained by the IBM 702 in the case of the solution of the differential equation (the formula is shown in Figure 6). There were executed approximately:

60,000 additions and subtractions
120,000 multiplications
20,000 divisions
50,000 logical decisions
10,000 figures printed with 7 digits each.

The printed problem solution is shown in Figure 7. The answer consists of the evaluation of eighteen basic series of p, q, r, s, t which are needed for the correct selection of the power series and of the evaluation of the amplitudes of the transient vibrations in the valve gear system for a great variety of power series and for a wide range of speed ratios. This work was necessary to determine the relationship between the selected cam profiles and the transient vibrations in the system in order to avoid valve failures.

In the meantime, many more problems have been analyzed and programmed. Some of them required only the translation of existing equations into machine language, others required an extensive analysis for several months. Many mathematical procedures and so-called "subroutines" have been employed for the problem solutions.

Step 1

$$y = h_r + \delta y + \phi y''$$

$$y = h(1 + \Sigma N_n^o x^n) + V\Sigma N_n^v x^n + A\Sigma N_n^a x^n.$$

$$y = h\Sigma\alpha N_n^o x^n + V\Sigma\alpha_n N_n^v x^n + A\Sigma\alpha_n N_n^a x^n$$

$$x = \theta/c;\ \alpha = \frac{n(n-1)}{x^2\, 10^4}$$

$$N_n^o = \frac{-pqrst/_n}{\pi(m-n)/_{(\omega-n)}} \qquad N_n^v = \frac{-\Sigma(n-6-\omega)100}{\pi(m-n)/_{(\omega-n)}}$$

$$N_n^a = \frac{1000}{\pi(m-n)/_{(\omega-n)}}$$

$$\int y/_h\, dx = 1 + \Sigma\frac{N_n^o}{n+1} + \pi\Sigma\frac{N_n^v}{n+1} + \beta\Sigma\frac{N_n^a}{n+1}$$

Step 2

$$S_c = 2m/\theta^{2m}\left(1 - \theta^2/2 + \ldots. (-1)^{m-1}\ \theta^{2m-3}/(2m-2)\right)$$

$$S_o = 2m-1/\theta^{2m-1}\left(\theta - \ldots\ldots (-1)^{m-1}\ \theta^{2m-3}/(2m-3)\right)$$

Figure 6

Equations of Cam Calculation

In Figure 8, a list of engineering automotive problems is given all of which have been successfully solved by the IBM 705. As far as possible, approximate time figures are also given.

In the following, some basic ideas about subroutines are presented. The significance of subroutines is based on the fact that they might be permanently stored on cards or tape and only the transfer addresses be known. They can then be used in the solution of any program involving subroutines. Subroutines might be preferably programmed in floating decimal point. But it is not always favorable to use floating decimal point mathematics in engineering problems. Therefore, either a conversion routine or a second subroutine procedure with fixed decimal point should be available.

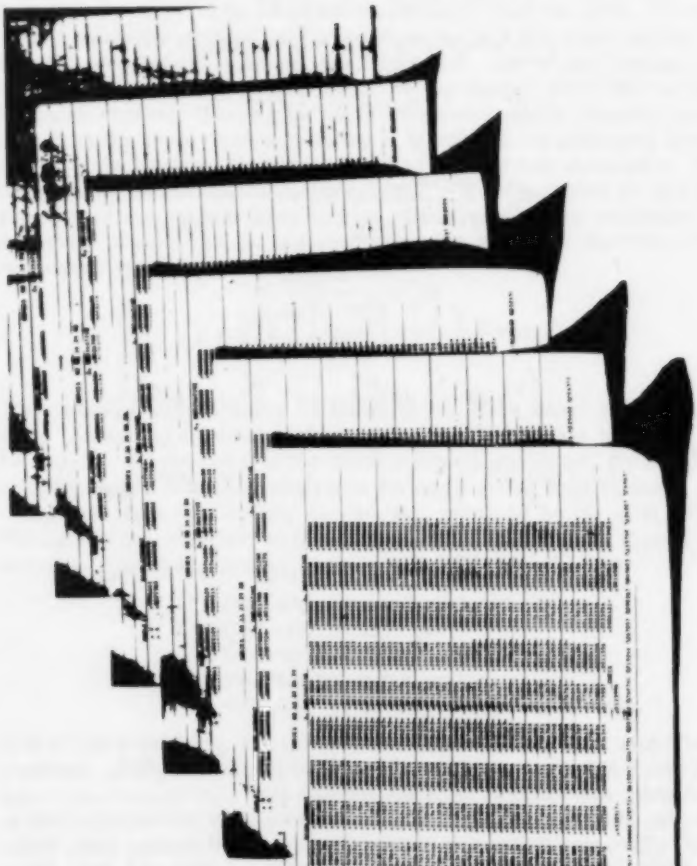


Figure 7

The Printed Problem Solution of the Differential Equation in Figure 6

Project	Desk Calc. Time Hours	IBM 705 Time Minutes
General Design Studies and Specific Design Solutions		
Torque Converter Design	40	1
Performance Studies	8-40	2-1
Cam Design	100	1.5
Pressure - Volume Diagram	25	.5
Rear Propeller Shaft Data	3	.06
Heat Regenerator for Gas Turbines	?	
Combustion Chamber Design	120	2
Fuel Economy		
Gear Designs (Spur, Helical, Bevel, Hypoids)		
Vibration Problems		
Truck Vibrations	40	1
Tractor-Trailer Vibrations (3-mass system)	50	1
Transient Vibrations in Valve Train	25	.5
Engine Balancing	?	
Stress Problems		
Bearing Loads	250	6
Spring Design	8	.2
Strength of Front & Rear Axle Shafts	8	.2
Bending Stresses in Frames	16	.5
Bending Stresses in Discs	40	1
Stresses in Rotating Discs of Non-uniform Thickness	40	1
Stresses in Gear Teeth (bending and compression)	100	2
Statistics		
Statistical Analysis of Data	4-40	.1-1
Curve Evaluation by Least Squares	?	1
Miscellaneous		
Subroutine Library	-	-
Calculation of 36 Polynomials	1 Year	40
Center of Gravity of Parts	4	.1
Drill Charts	8	.2
Load Distribution in Trucks	-	-

Figure 8

Engineering automotive problems solved by the IBM 705

Extracting Roots

It appears that the interation method by Newton

$$x_1 = x_0 - \frac{f(x)}{f'(x)}$$

is the best for extracting roots of any order. Let

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots\ldots + a_1 x + a_0$$

be the polynomial equation of the n^{th} order, and thus the first derivative is

$$f'(x) = a_n n x^{n-1} + a_{n-1}(n-1)x^{n-2} + \ldots\ldots + a_0 .$$

Inserting f(x) and f'(x) in the iteration formula results in

$$x_1 = \frac{(n-1)a_n x_0 + (n-2)a_{n-1}x_0^{n-1} + \ldots\ldots + a_2 x_0^2 - a_0}{n a_n x_0^{n-1} + (n-1)a_{n-1}x_0^{n-2} + \ldots\ldots + 2a_2 x_0 + a_1} .$$

In order to obtain the form

$$x = \sqrt[n]{u}$$

the constants in the equation for f(x) are

$$a_n = 1$$
$$a_{n-1} \ldots\ a_1 = 0$$
$$a_0 = -u .$$

Then, the interation formula becomes

$$x_1 = \frac{1}{n}\left[(n-1)x_0 + \frac{u}{x_0^{n-1}}\right]$$

where u = radical, x_0 = first approximation, x_1 = second approximation, n = the order of required root.

This formula allows the extraction of roots of any order and can be programmed by the use of 39 instructions. Choosing n = 2, the known formula for extracting square roots is obtained:

$$x_1 = \frac{1}{2}\left[x_0 + \frac{u}{x_0}\right] .$$

Evaluation of Power Series

The evaluation of exponential, logarithm and trigonometric functions can be made by means of power series. All of the power series in question have the same pattern of increasing terms:

$$f(x) = C \pm C\frac{B}{E_1} \mp \frac{CB}{E_1} \cdot \frac{B}{E_2} \pm \frac{CB^2}{E_1E_2} \cdot \frac{B}{E_3} \mp \ldots\ldots$$

where C = first term of the series, B = factor by which the following term is increased, E_n = factor by which the following term is divided. This factor can have the form A or A(A-1), where A is a variable, depending upon H.

The factors for some series are given below:

	A	B	C	E	H
e^x	001	x	1	A	001
sin hx	003	x^2	x	A(A-1)	002
ln a *)	003	x^2	x	A	002
sin x	003	x^2	x	A(A-1)	002
$\sin^{-1}x$	003	x^2	x	A(A-1)	002

*) $x = \frac{a-1}{a+1}$

The number of terms to be used for one value depends upon the desired accuracy. If the answer must be accurate to seven digits after the decimal point, then n terms have to be used so that the (n + 1)th term does not add any significant digits to the answer obtained by the summation of n terms.

Trial and Error Procedure

It appears that there is one general method which covers all possible trial and error procedures. Examples in which trial and error procedures are involved are polynomials of higher than the second order, transcendental functions, and so forth. Let us assume that the number .1638 is the right answer of a transcendental function. The awkward way to approach this number would be to insert the initial value of .0001 and increase this value by .0001. In this way, the computer had to execute 1,638 loops in order to obtain the right answer. The other possibility is the following:

Insert .1000 and increase by .1000. This new value .2000 is too high and has to be decreased. This can be done by shifting the

decimal point to the left and subtracting. Therefore, the new value of .1900 is obtained. This procedure has to be repeated until the result is found.

.1900
.1800
.1700
.1600
.1610
.1620
.1630
.1640
.1639
.1638

It appears that this method requires only twelve loops instead of 1,638 loops.

Differential and Integral Calculus

Every differentiation or integration consists of two parts:

1. The mathematical solution,
2. The numerical evaluation of the solution.

A digital computer cannot help in part 1 but it can help in part 2. The numerical treatment can be broken down into equations or series by programming for a digital computer.

In the last four examples of subroutines, it has been shown that many mathematical operations, in addition to the four arithmetic operations, can be performed successfully by the automatic computers.

SUMMARY

It has been shown that the increased complexity of engineering problems requires new, powerful labor-saving devices in order to reduce the routine efforts of highly skilled technical personnel and in order to make possible the solution of many problems not efficiently soluble at the present time.

Based on experience gained in actual problems, it is believed that the new large-scale electronic computers will meet these requirements.

REFERENCES

1. R. F. Clippinger, B. Dimsdale, and T. H. Levin, "Automatic Digital Computers in Industrial Research I - V," *Journal of the Society for Industrial and Applied Mathematics*, 1 - 2, II, 1 - 3.

2. C. B. Tompkins and T. H. Wakelin, *High Speed Computing Devices*, New York: McGraw-Hill Book Company Inc., 1950.

3. W. V. Wilkes, D. J. Wheller, S. Gill, *The Preparation of Programs of an Electronic Computer*, Cambridge, England: University Mathematical Laboratory, 1951.

4. *Annals of the Computation Laboratory of Harvard University*, Vols. 1, 16, 18, 19, 20, 22, 26, 27.

5. IBM, *Proceedings of Computation Seminar* and *Technical News Letter*.

4

Tire Vulcanization—An Application of Unsteady Heat Transfer

By Samuel R. Schaffer
United States Rubber Co.
Detroit

1. INTRODUCTION

A. Vulcanization

The change which took place when Goodyear accidently caused rubber to be combined with sulfur is described in today's terms as a cross-linking reaction, in which the long chain rubber molecules are tied together by means of sulfur bonds to form a 3-dimensional network. A simplified picture of cross-linked or cured rubber is shown in Figure 1 below. Each of the chains consists of thousands of repeated molecular units. Goodyear termed this process vulcanization after Vulcan, the Roman god of fire, because heat was necessary to carry out the reaction.

Figure 1
Pictorial Representation of Cross Linked Rubber

Natural rubber and most of the synthetic rubbers used in large volume are still vulcanized by reacting with sulfur at elevated temperatures.

The most important consequence of vulcanization is that the rubber is converted from a plastic to an elastic material, i.e., the strong chemical forces at the sulfur bonds restrain the chains from sliding past each other. Therefore, cured rubber which is deformed under load will retract with nearly complete recovery when the load is released.

B. Effect of State of Cure on Properties of Vulcanized Rubber

The degree or extent to which this reaction is carried is known as the "state of cure." The state of cure is given by $\int Rd\,\theta$, where the rate R depends on temperature and generally varies in time.

As the curing reaction proceeds and the state of cure increases, the physical properties of the rubber compound, such as tensile strength, tear resistance, abrasion resistance, increase rapidly, approach a maximum and then recede somewhat more slowly. A generalized curve of properties vs. cure will look like the curve in Figure 2.

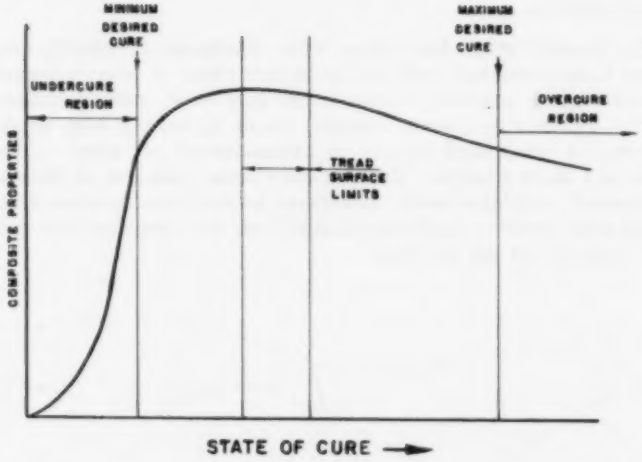


Figure 2

Generalized Curve of Compound Properties Vs. State of Cure

Since the performance of rubber items will depend on one or more of these properties, it is necessary that the state of cure be controlled within the range that provides optimum properties. In commercial vulcanization of most rubber products, the temperature of the unit is raised by applying heat to the outer surfaces and depending on the conduction of heat through the rubber to raise the temperature at the center. The outer portions are therefore subject to higher temperatures than the interior regions, and the curing reaction proceeds faster at the outside than the inside.

It is desirable to adjust the vulcanization time and temperature so that in the region of least cure, the cure has proceeded

sufficiently far so that it is out of the undercure range, while in the region of maximum cure it has not reached the overcure range.

This ratio of maximum to minimum acceptable cure might be 3:1 or 5:1, depending on the compound. The maximum and minimum values establish the first set of cure limits that must be met.

In the case of tires, a narrower set of cure limits is required for the outer surface of the tread, as the resistance to cutting and chipping falls off rapidly with higher cure, while tread wear falls off with lower cure.

For tires a third limit is imposed by the fact that either steam or hot water under pressure and contained within a rubber curing bag is used as the source of internalheat, and that some time before the end of the cure cycle it is necessary to exhaust the internal pressure down to zero so that the press can be safely opened. As the bag pressure is reduced, the water vapor and air which were trapped in the interior of the tire during the earlier assembly operations try to expand and blow the tire apart. The third condition then is that the cure has proceeded far enough at the time of bag exhaust so that the stocks have sufficient strength to withstand the internal forces due to trapped gases.

2. ONE DIMENSIONAL TRANSIENT FLOW

A. General Case

We have imposed a number of limits on the state of cure and have indicated that state of cure = $\int R d\theta$ where the rate R depends on temperature. For the rubber compounds used in tires, we find that the rate of cure approximately doubles for a 16°F increase in temperature. It is necessary, then, to know the temperatures that exist in the tire during the cure cycle in order to set up a cycle such that the cure falls within the specified limits.

The physical system we are dealing with consists of a mold, tire and bag, with the outer mold face heated by steam and the inner bag surface heated by steam or hot water. The heat flow is from outside toward the center and can be approximated as a problem in one dimensional flow. To examine the conditions for heat flow in one direction, consider an element of thickness x and area A. The heat balance for this element is:

Heat flow in - heat flow out + heat generated = heat stored.

The curing reaction is slightly exothermic, but for soft rubber compounds the heat generated is small and can be ignored. (For hard rubber compounds much larger amounts of sulfur are used and the heat generated by the reaction becomes important).

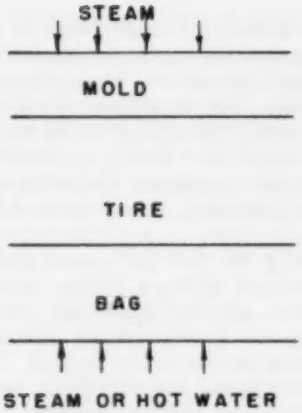


Figure 3
Application of Heat During Tire Cure

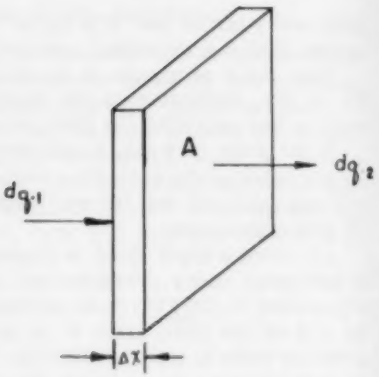


Figure 4
One Dimensional Heat Conduction

The rate of heat flow by conduction is given by the equation:

$$\frac{dq}{d\theta} = -kA \frac{\partial t}{\partial x}$$

where θ = time,
k = thermal conductivity.

The minus sign is required because the direction of positive heat flow is from higher to lower temperatures. Then writing the heat balance as a rate equation, we get:

$$-kA \left(\frac{\partial t}{\partial x}\right)_{x_1} + kA \left(\frac{\partial t}{\partial x}\right)_{x_1 + \Delta x} = A(\Delta x) c\rho \frac{\partial t}{\partial \theta} .$$

This reduces to:

$$\frac{k}{c\rho} \frac{\partial^2 t}{\partial x^2} = \frac{\partial t}{\partial \theta} \tag{1}$$

which is Fourier's equation for heat conduction in one direction.

The term $\frac{k}{c\rho}$ is called diffusivity and is usually designated by α.

The solutions to this equation for many regularly shaped solids, such as slabs, cylinders, spheres and bricks are given in a number of standard texts.

B. Solution of One Dimensional Transient Flow Equation for Infinite Slab

For an infinite slab of thickness L, initially at uniform temperature t_i, then at $\theta = 0$, suddenly exposed to constant surface temperature t_s on both faces, the solution is:

$$(2) \quad t_{(x,\theta)} = t_s + (t_i - t_s)\frac{4}{\pi}\sum_{n=1,3,5\cdots}^{\infty}\frac{1}{n}\operatorname{Sin}\frac{n\pi x}{L}\, e^{-n^2\pi^2\alpha\,\theta/L^2}$$

This is a rapidly converging series and for most cases only the first term need be considered (after a relatively short time).

For the temperature at the midplane, $x = L/2$, this becomes:

$$(3) \quad t_{(c,\theta)} = t_s + (t_i - t_s)\frac{4}{\pi}\, e^{-\pi^2\alpha\,\theta/L^2}$$

From this we can see that:

1. The time required to reach a given temperature is proportional to L^2.
2. The time required to reach a given temperature is inversely proportional to α.

C. Establishment for Cure Cycles from Temperature Data

The problem presented in curing tires can best be appreciated if we consider that rubber has the lowest value of α that is reported in Ingersoll and Zobel's *Heat Conduction.*

Compounded rubber: $\alpha = .005$ Ft^2/hr

Values of α for other typical insulating materials:

.006 (silica aerogel) to .010 (asbestos).

Values of α for metals: aluminum 3.3
steel 0.48

This means that if a set of slabs made of these materials all initially at the same temperature were exposed to a new surface temperature, the rubber slab would have the slowest temperature rise at interior points. The next figure shows some temperatures measured during the cure of a 9.00-20 tire. This is the size that might be used on a city bus.

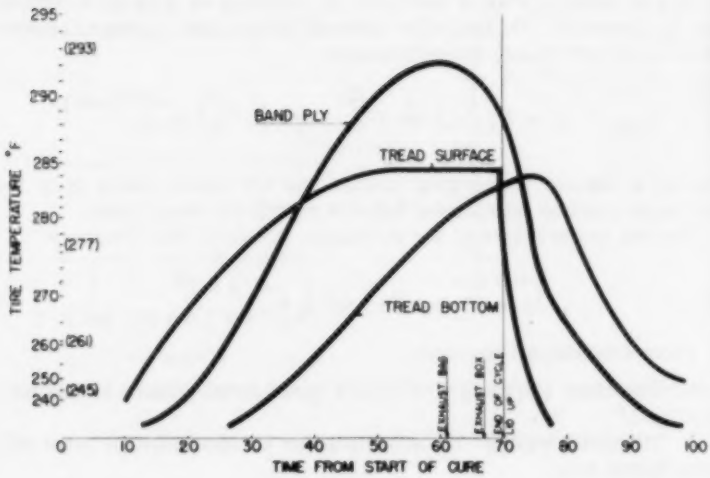


Figure 5

Temperature Rise in 9.00-20 8/10 Fleetway.
70' McNeil U.V. Cure Cycle. 286°F External.
310°F Circulating Hot Water

Note that the temperatures are plotted on a scale that expands exponentially rather than a linear scale. The temperature scale has been designed so that the vertical distances are proportional to the rate of cure rather than the temperature. Conversion from temperature to rate of cure is based on the fact that the temperature coefficient is slightly greater than 2 so that the rate doubles for each 16°F temperature increase.

The area under the rate-time curves gives state of cure. Using temperatures measured by means of thermocouples, the state of cure values for the surfaces and interior are calculated and checked against the specified state of cure limits.

Note that in the actual case of a tire cure the temperature at the surface does not increase abruptly with time, so that the ideal boundary conditions set up for equations (2) and (3) are not met. How then can these temperature measurements be used to estimate or calculate the temperatures for a new set of conditions?

3. ESTIMATE TEMPERATURES DURING CURE CYCLES

Let us consider the system consisting of the mold, tire and bag. Each of these has a different thickness and conductivity. However, we can think of the mold and bag as being replaced by slabs of rubber having the same thermal properties as the tire. This system is shown in Figure 6. The outer faces of the mold and bag, which are exposed to steam and hot water, respectively, come much closer to meeting the assumption of constant surface temperatures than do the tire surfaces.

d_m, d_b = ORIGINAL MOLD AND BAG THICKNESS

α_m, α_b = " " " " DIFFUSIVITY

d_{m_1}, d_{b_1} = REDUCED " " " THICKNESS CORRESPONDING TO TIRE DIFFUSIVITY, α_t

TOTAL THICKNESS OF EQUIVALENT HEAT TRANSFER SYSTEM

$$L = d_{m_1} + d_t + d_{b_1}$$

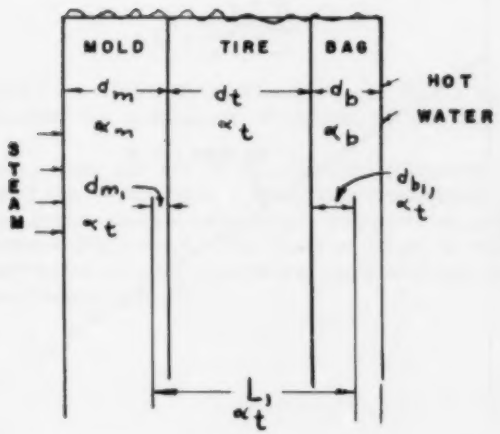


Figure 6

Replacement of Mold and Bag With Slabs Having Diffusivity of Tire Material

Consider this system as a single slab, for which we wish to know temperatures at various depths. The general solution when both surface temperatures are equal is still given by equation (2). Using only the first term of the series, we can simplify equation (2) by substituting

$$A = \frac{4}{\pi}, \quad A_x = A \sin\frac{\pi x}{L}, \quad B_x = \left(\frac{t_i}{t_s} - 1\right) A_x, \quad \text{and} \quad C = \frac{\pi^2 \alpha}{L^2}$$

$$t_{(x,\theta)} = t_s + (t_i - t_s)\, A \sin\frac{\pi x}{L}\, e^{-c\theta} \tag{4}$$

$$= t_s\left[1 + \left(\frac{t_i}{t_s} - 1\right) A_x e^{-c\theta}\right] = t_s\left(1 - B_x e^{-c\theta}\right).$$

We can write an equation of this form for any depth in the slab including the tire surfaces and center.

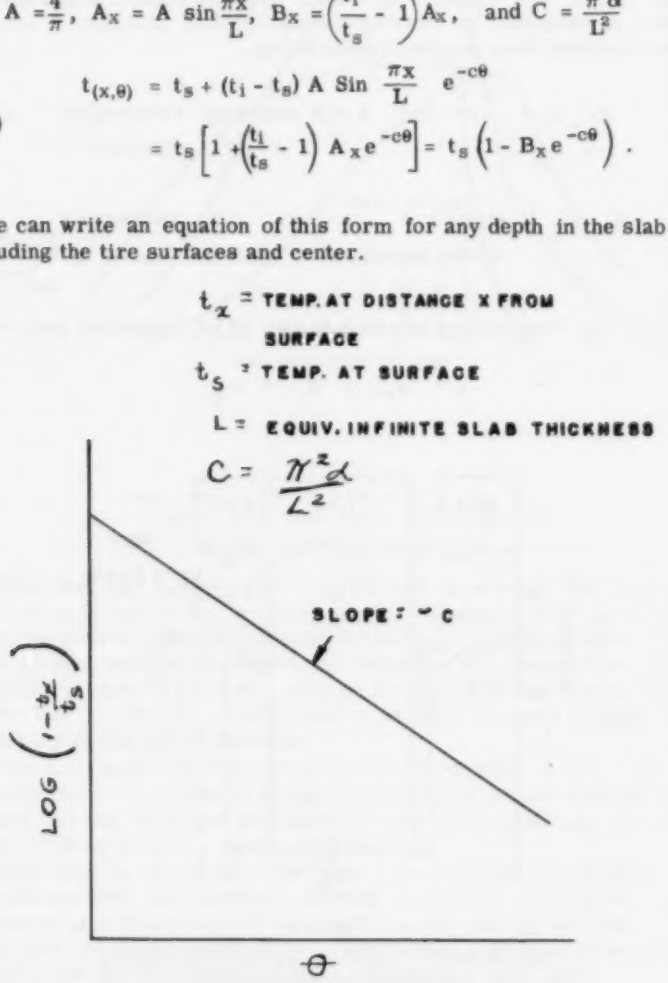

t_x = TEMP. AT DISTANCE X FROM SURFACE
t_s = TEMP. AT SURFACE
L = EQUIV. INFINITE SLAB THICKNESS
$C = \frac{\pi^2 \alpha}{L^2}$
SLOPE = $-C$
$\log\left(1 - \frac{t_x}{t_s}\right)$
θ

Figure 7

Determination of Equivalent Infinite Slab Thickness From Tire Temperature Data

Rearranging and taking logs gives:

$$\text{Log}\left(1 - \frac{t_x}{t_s}\right) = \text{Log } B_x - c\theta \,. \tag{5}$$

Equation (5) shows that for constant mold steam and bag water temperature a plot of a simple function of the temperature at any thickness, x, within the tire should be linear in time.

By differentiating equation (4) with respect to time, we get:

$$\frac{dt_{x,\theta}}{d\theta} = c\,(t_s - t_{x,\theta}) \,. \tag{6}$$

The temperature difference $t_s - t_{x,\theta}$ can be thought of as the driving force tending to change the temperature at depth x; at any time θ the rate of change of temperature at x is proportional to this driving force. This equation for temperature change in unsteady state conduction is analogous to the equation for heat flow in steady state conduction, which states that the rate of heat flow is proportional to the driving force of temperature difference.

With temperature measurements available at various depths in a tire under one set of boundary conditions, equations (5) and (6) can be used to estimate new temperature curves and new state of cure values corresponding to new boundary conditions. Further, since

$$C = \frac{\pi^2 \alpha}{L^2}$$

and α is known or readily determined, we can find a value for L, which represents the equivalent thickness of the heat transfer system.

In this way, from one set of experimental temperature values, we can predict the results under many other conditions, including new boundary temperatures, new initial temperatures, and changes in the equivalent thickness L. The final estimate is usually confirmed by additional direct temperature measurements using thermocouples embedded in the tire.

5

Some
Valid
from

By Le
Gene

Of
most
given
tions
usuall
hand
reaso

A
a rea
some
the g
value
value
out of
tion f

Exam

Le
bearin
first
bunch

Le
ulatio
validi
from
betwe
based
choos
first
above

5

Some Non-Parametric Methods of Testing the Validity of Assuming that a Sample Comes from a Given Population

By Leonard G. Johnson, Research Staff
General Motors Corporation, Detroit

INTRODUCTION

Of all questions which arise in sampling statistics, one of the most important is whether or not it is reasonable to assume that a given sample comes from some definite population. Marked deviations of a set of one or more sample points from a population line usually create a suspicion as to whether the particular sample at hand really is a member of the family of samples which have a reasonable chance of being randomly picked from the population.

A convenient method of determining whether or not a sample has a reasonable chance of coming out of a given population is to choose some meaningful "statistic," i.e., some function of one or more of the given sample points, and then determine where the numerical value of the statistic falls with respect to the complete range of values of the statistic for all possible samples of the same size taken out of the population. This entails the determination of the distribution function of the particular statistic chosen as a criterion.

Example I - A Test Based on the Two Lowest Values

Let us assume that a fatigue test has been run on some ball bearings, and we observe that there is a wide separation between the first and second failures, and that the succeeding failures are bunched close together, as shown in Figure 1.

Let us further suppose that the line AB is claimed to be the population line for this sample. We then raise the question as to the validity of assuming that such a sample could very readily come from the population represented by line AB. The wide separation between the first and second failures suggests the use of a statistic based on these two failures alone as a criterion. Let us therefore choose the non-parametric statistic given by the probability that the first failure lies below rank r_1, and that the second failure lies above rank r_2.

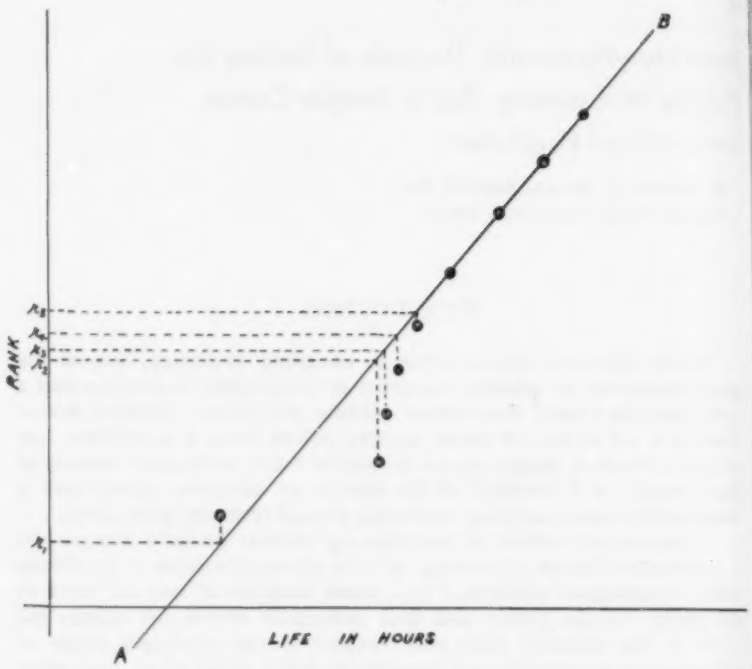


Figure 1

In order to obtain a formula for this statistic, we draw the following sketch on the rank scale, represented by the letter z:

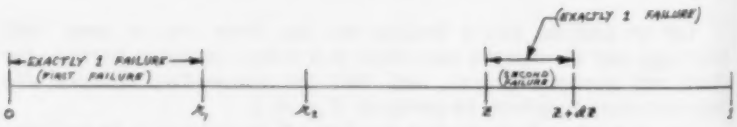


Figure 2

The probability that the second failure is located between the ranks z and z + dz is $g_2(z)dz$ where $g_2(z)$ is the probability density function of the rank of the second failure.

The probability that the first failure lies below rank r_1, *if* the second failure is located between the ranks z and z + dz, is $\frac{r_1}{z}$.

Therefore, the joint probability that the first failure lies below rank r_1, *and* that the second failure is located between ranks z and

$z + dz$ is the product, $\left(\frac{r_1}{z}\right)g_2(z)dz$. But, the possible positions for z are anywhere from r_2 to 1. Hence, the probability that the first failure lies below rank r_1, *and* that the second failure lies above rank r_2 is

$$P_n = \int_{r_2}^{1} \left(\frac{r_1}{z}\right)g_2(z)dz .$$

It has been shown that for a sample of size n

$$g_2(z) = n(n-1)\, z\, (1-z)^{n-2} .$$

(See *Industrial Mathematics*, Vol. 2, p. 3, Formula (A).)

$$\therefore\ P_n = \int_{r_2}^{1} \left(\frac{r_1}{z}\right)n(n-1)\, z\, (1-z)^{n-2}dz = n(n-1)r_1\int_{r_2}^{1} (1-z)^{n-2}dz$$

$$= \left[-n(n-1)r_1 \frac{(1-z)^{n-1}}{n-1}\right]_{r_2}^{1} = nr_1(1-r_2)^{n-1} ,$$

which is the desired statistic to be used as a criterion.

Suppose now, for the sample in Figure 1, that n = 50, $r_1 = .011$, and $r_2 = .074$. Then, for such a sample, $P = 50\,(.011)\,(1-.074)^{49}$
$= .0127.$

This tells us that, for a sample of 50 failures, the probability that the first failure lies below rank .011, *and* that the second failure lies above rank .074 is .0127. That is, there are about 127 chances in 10,000 that a sample of 50 would have the first failure below rank .011 and the second failure above rank .074. One might decide from this that the sample does not come from the population represented by the line AB. But, how often would such a decision be wrong? This is a more fundamental question which should be answered before any decision is made at all.

We can not say that the probability that this decision is wrong is .0127, because of the fact that even for the case in which the first and second points fall exactly on the line AB at median ranks, i.e., when $r_1 = .0138$ and $r_2 = .0336$, the value of the statistic is $P_n = 50\,(.0138)\,(1-.0336)^{49} = .1293$, and we know that there are much more than 1293 chances in 10,000 of being wrong when we reject the

hypothesis that such a "perfect sample" comes from the population represented by the line AB. It therefore becomes necessary to determine the distribution function of the statistic $P_n = nr_1(1-r_2)^{n-1}$ for all possible samples of size n coming out of the given population. To do this, we proceed as follows:

The probability density function of r_2 is

$$f(r_2) = n(n-1)r_2\,(1-r_2)^{n-2}\ .$$

Let $\quad y = (1-r_2)^{n-1}\ ,\ \text{or}\ \ r_2 = 1 - y^{\frac{1}{n-1}}$

$$dr_2 = -\left(\frac{y^{\frac{2-n}{n-1}}}{n-1}\right)dy$$

$$\therefore\ \left|f(r_2)dr_2\right| = \left|n(n-1)\,(1-y^{\frac{1}{n-1}})\,y^{\frac{n-2}{n-1}}\left(-\frac{y^{\frac{2-n}{n-1}}}{n-1}\right)dy\right| = n\,(1-y^{\frac{1}{n-1}})dy.$$

Thus, the probability density function of the quantity $y = (1-r_2)^{n-1}$ is

$$f(y) = n(1 - y^{\frac{1}{n-1}})\ .$$

We can now determine the distribution function of the product $q = r_1(1-r_2)^{n-1} = xy$ as follows: (x represents r_1, and y represents $(1-r_2)^{n-1}$.)

$$\text{Prob.}\ (a \leqslant y \leqslant a + da) = n\left(1 - a^{\frac{1}{n-1}}\right)da$$

$$\text{Prob.}\left[(x < r_1) \subset (a \leqslant y \leqslant a + da)\right] = \frac{r_1}{r_2} = \frac{r_1}{\left(1-y^{\frac{1}{n-1}}\right)_{y=a}} = \frac{r_1}{1-a^{\frac{1}{n-1}}}$$

$$\text{Prob.}\left[\left(x < \frac{q_0}{a}\right) \subset (a \leqslant y \leqslant a + da)\right] = \frac{q_0}{a\left(1-a^{\frac{1}{n-1}}\right)}$$

$\therefore$ By multiplication:

$$\text{Prob.}\left[\left(x < \frac{q_0}{a}\right) \cdot (a \leqslant y \leqslant a + da)\right] = \frac{q_0}{a\left(1-a^{\frac{1}{n-1}}\right)} \cdot n\left(1-a^{\frac{1}{n-1}}\right)da$$

$$= nq_0\,\frac{da}{a}\ .$$

Another way of writing this last probability is

$$\text{Prob.}\left[(xy < q_0) \cdot (a \leqslant y \leqslant a + da)\right] = nq_0\,\frac{da}{a}\ .$$

In order to determine the total probability that $xy < q_o$, we must integrate $nq_o \frac{da}{a}$ between the lower and upper limits of $\underline{a}$.

Thus, Prob. $_{pp}(xy < q_o) = \int_{a_1}^{a_2} nq_o \frac{da}{a} = nq_o \operatorname{Log}_e\left(\frac{a_2}{a_1}\right)$,

where the subscript pp denotes the principal part of the probability, and where $\underline{a_1}$ and $\underline{a_2}$ are respectively the lower and upper limits of $\underline{a}$. These limits are determined from the fact that

$$\frac{q_o}{a\left(1 - a^{\frac{1}{n-1}}\right)} \leq 1.$$

Therefore, in the principal part, $\underline{a}$ may range over all values for which the inequality $a\left(1-a^{\frac{1}{n-1}}\right) \geq q_o$ holds. Thus, it can be seen that the range of integration on $\underline{a}$ is a function of q_o.

For the particular case which we are considering, we have

$$n = 50, \text{ and } q_o = \frac{P_h}{50} = \frac{.0127}{50} = .000254 .$$

Hence, the range of integration is determined by the inequality

$$a\left(1 - a^{\frac{1}{49}}\right) \geq .000254.$$

This inequality is satisfied by all values of $\underline{a}$ in the interval $.002157 \leq a \leq .987474$. Thus, for this case, $a_1 = .002157$ and $a_2 = .987474$.

$$\therefore \text{ Prob. }_{pp}\left[r_1(1-r_2)^{n-1} < .000254\right] = 50(.000254) \operatorname{Log}_e\left(\frac{.987474}{.002157}\right)$$

$$\text{or, Prob. }_{pp}(P_n < .0127) = .0127 \operatorname{Log}_e\left(\frac{.987474}{.002157}\right) = .0778 .$$

To this principal part we must add the complementary part of the probability, which is equal to

$$\int_0^{a_1} f(y)dy + \int_{a_2}^1 f(y)dy = 1 + F(a_1) - F(a_2),$$

where $$F(y) = \int_0^y n\,(1-y^{\frac{1}{n-1}})\,dy = ny - (n-1)y^{\frac{n}{n-1}} .$$

Hence, the complementary part of the probability is

$$\begin{aligned} \text{Prob.}_{cp}(P_n < .0127) &= 1 + F(.002157) - F(.987474) \\ &= 1 + 50(.002157) - 49(.002157)^{\frac{50}{49}} \\ &\quad - 50(.987474) + 49(.987474)^{\frac{50}{49}} \\ &= .0147 . \end{aligned}$$

$$\therefore \quad \text{Prob.}\ (P_n < .0127) = .0778 + .0147 = .0925.$$

This last number represents the probability of our being wrong when we reject the hypothesis that the given sample comes from the population represented by the line AB on the basis that the value of P_n is as small as .0127.

Example II - Another Test Based on the Two Lowest Values

Another statistic which might be used as a criterion and which involves only the two lowest values is the probability of obtaining a separation in rank between the first and second failures which is at least as great as $h = r_2 - r_1$.

The formula for this probability is obtained as follows:

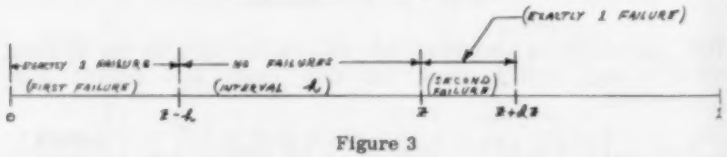


Figure 3

The probability that the second failure is located between ranks z and z + dz is $g_2(z)dz$, where $g_2(z)$ is the probability density function of the rank of the second failure.

The probability that the first failure lies below rank (z-h), *if* the second failure lies between ranks z and z + dz is $\frac{z-h}{z}$.

Therefore, the joint probability that the first failure lies at least h rank units below the second, *and* that the second is located between the ranks z and z + dz is the product $\left(\frac{z-h}{z}\right) g_2(z)dz$.

The possible positions for z are anywhere from h to 1. Hence, the probability of a separation of at least h in rank between the first and second failures is

$$Q_n = \int_h^1 \left(\frac{z-h}{z}\right) g_z(z)\,dz = \int_h^1 \left(\frac{z-h}{z}\right) \cdot n(n-1)\, z\, (1-z)^{n-2}\, dz$$

$$= n(n-1)\int_h^1 (z-h)(1-z)^{n-2}\, dz .$$

Integrating by parts: $u = z - h$ $\qquad dv = (1-z)^{n-2}\, dz$

$$du = dz \qquad v = -\frac{(1-z)^{n-1}}{n-1}$$

$$Q_n = n(n-1)\left\{\left[-\frac{(z-h)(1-z)^{n-1}}{n-1}\right]_h^1 + \frac{1}{n-1}\int_h^1 (1-z)^{n-1}\, dz\right\}$$

$$= \left[-n(n-1)\cdot\frac{1}{n-1}\cdot\frac{(1-z)^n}{n}\right]_h^1 = (1-h)^n, \text{ which}$$

is the desired formula.

For the special case considered previously, we had $r_1 = .011$ and $r_2 = .074$ with $n = 50$. Hence, $h = r_2 - r_1 = .063$, and $Q_n = (1-h)^n = (1-.063)^{50} = (.937)^{50} = .0386$. Thus, the value of the statistic is .0386, which represents the probability of having at least .063 units of separation in rank between the first and second failures. Now if we reject the hypothesis that the sample comes from the population represented by line AB of Figure 1 on the basis of $Q_n = .0386$, what is the probability that we are wrong? We answer this question as follows:

The probability density function of h is

$$f(h) = n(1-h)^{n-1} .$$

$$\text{put} \quad y = (1-h)^n , \quad \text{or} \quad h = 1 - y^{\frac{1}{n}}$$

$$dh = -\frac{y^{\frac{1-n}{n}}}{n}\, dy$$

$$\therefore \left| f(h)dh \right| = \left| n\, y^{\frac{n-1}{n}} \left(-\frac{y^{\frac{1-n}{n}}}{n}\, dy \right) \right| = dy .$$

$\therefore$ The probability density function of the quantity $y = (1-h)^n$ is

$$f(y) \equiv 1 \qquad (o \leqslant y \leqslant 1)$$

$\therefore$ The cumulative distribution function of y is

$$F(y) = \int_o^y f(y)dy = y$$

$$\therefore \quad \text{Prob.} \quad (y < Q_n) = Q_n .$$

Thus, Prob. $(y < .0386) = .0386$, and we see that with this criterion the probability of a wrong decision is the same as the numerical value of the statistic. Since the probability of a wrong decision is less with this criterion, it follows that this criterion is superior to the first one considered. This, of course, need not be true for all samples, just because it happens to be true in this particular case.

Example III - A Test Based on the First Three Failures

As a third example of a criterion we shall consider the probability that the first failure lies below rank r_1, and that the second and third failures are confined to a k - band whose upper boundary is at or above rank z.

We derive the formula for this probability as follows:

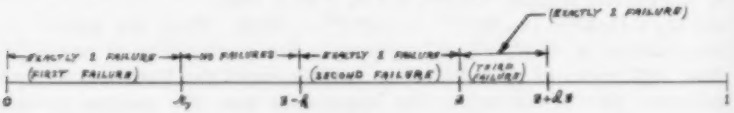


Figure 4

$$\text{Prob.} \left[\begin{array}{l} \text{(3rd Failure between z and z + dz) and (exactly 1 failure} \\ \text{below } r_1\text{) and (exactly 1 failure between (z-k) and z)} \end{array} \right]$$

$$= g_3(z)dz \cdot \frac{2r_1}{z} \cdot \frac{k}{z} = \frac{2r_1 k}{z^2}\, g_3(z)\, dz$$

where $g_3(z)$ is the probability density function of the rank of the third failure. By integrating $\frac{2r_1 k}{z^2}\, g_3(z)\, dz$ over the range $z = z$ to $z = 1$ we obtain the desired probability as

$$R_n = 2\int_z^1 \frac{r_1 k}{z^2}\, g_3(z)\, dz .$$

Now, $$g_3(z) = \frac{n(n-1)(n-2)}{2} z^2 (1-z)^{n-3}$$

(See *Industrial Mathematics*, Vol. 2, p. 3, Formula (A).)

$$\therefore \quad R_n = \frac{2r_1 kn(n-1)(n-2)}{2} \int_z^1 (1-z)^{n-3} \, dz$$

$$= \frac{2r_1 kn(n-1)(n-2)}{2} \left[-\frac{(1-z)^{n-2}}{n-2} \right]_z^1$$

$$= r_1 kn(n-1)(1-z)^{n-2} .$$

For the sample plotted in Figure 1 we have $n = 50$, $r_1 = .011$, $k = .013$, $z = .087$.

$$\therefore \quad R_n = (.011)(.013)(50)(49)(.913)^{48} = .0044.$$

If, now, on the basis of this value of R, we reject the hypothesis that this sample comes from the population represented by line AB, what is the probability that we are wrong? To answer this question we must determine the distribution function of the product $\xi = r_1(1-z)^{n-2}$. We do this as follows:

The probability density function of z is

$$f(z) = \frac{n(n-1)(n-2)}{2} z^2 (1-z)^{n-3} .$$

Let $y = (1-z)^{n-2}$, or $z = 1 - y^{\frac{1}{n-2}}$

$$dz = -\frac{y^{\frac{3-n}{n-2}}}{n-2} dy .$$

$$\therefore \left| f(z)dz \right| = \left| \frac{n(n-1)(n-2)}{2} \left(1 - y^{\frac{1}{n-2}}\right)^2 \cdot y^{\frac{n-3}{n-2}} \left(-\frac{y^{\frac{3-n}{n-2}}}{n-2} dy \right) \right|$$

$$= \frac{n(n-1)}{2} \left(1 - y^{\frac{1}{n-2}}\right)^2 dy$$

Thus, the probability density function of the quantity $y = (1-z)^{n-2}$ is

$$f(y) = \frac{n(n-1)}{2} \left(1 - y^{\frac{1}{n-2}}\right)^2 .$$

We can now determine the distribution function of the product $\xi = r_1(1-z)^{n-2} = xy$ as follows:

(x represents r_1, and y represents $(1-z)^{n-2}$.)

$$\text{Prob.}\ (a \leqslant y \leqslant a + da) = \frac{n(n-1)}{2}\left(1-a^{\frac{1}{n-2}}\right)^2 da$$

$$\text{Prob.}\left[(x < r_1) \subset (a \leqslant y \leqslant a + da)\right] = \frac{r_1}{z} = \left[\frac{r_1}{1-y^{\frac{1}{n-2}}}\right]_{y=a} = \frac{r_1}{1-a^{\frac{1}{n-2}}}.$$

$$\text{Prob.}\left[\left(x < \frac{\xi_0}{a}\right) \subset (a \leqslant y \leqslant a + da)\right] = \frac{\xi_0}{a\left(1-a^{\frac{1}{n-2}}\right)}$$

$$\therefore \text{Prob.}\left[\left(x < \frac{\xi_0}{a}\right) \cdot (a \leqslant y \leqslant a + da)\right] = \frac{n(n-1)}{2}\left(1-a^{\frac{1}{n-2}}\right)^2 \frac{\xi_0}{a\left(1-a^{\frac{1}{n-2}}\right)} da$$

$$= \frac{n(n-1)}{2} \cdot \frac{\xi_0}{a} \cdot \left(1-a^{\frac{1}{n-2}}\right) da .$$

The last probability may also be put into the form

$$\text{Prob.}\left[\ xy < \xi_0) \cdot (a \leqslant y \leqslant a + da)\right] = \frac{n(n-1)}{2}\left(\frac{\xi_0}{a}\right)\left(1-a^{\frac{1}{n-2}}\right) da .$$

$$\therefore \text{Prob.}_{pp}(xy < \xi_0) = \frac{n(n-1)\xi_0}{2}\int_{a_1}^{a_2}\left(\frac{1}{a} - a^{\frac{2-n}{n-2}}\right) da$$

$$= \frac{n(n-1)\xi_0}{2}\left[\text{Log}_e\left(\frac{a_2}{a_1}\right) + (n-2)\left(a_1^{\frac{1}{n-2}} - a_2^{\frac{1}{n-2}}\right)\right].$$

Again, the subscript pp denotes the principal part of the probability, while a_1 and a_2 are the limits on a between which the following inequality holds:

$$\frac{\xi_0}{a\left(1 - a^{\frac{1}{n-2}}\right)} \leqslant 1$$

$$\text{or,} \qquad a\left(1 - a^{\frac{1}{n-2}}\right) \geqslant \xi_0$$

In our particular example,

$$R_n = .0044, \quad \text{or} \quad \xi_0 = \frac{R_n}{kn(n-1)} = \frac{.0044}{(.013)(50)(49)} = .000138.$$

$\therefore$ The limits on a for the principal part are to be determined from the inequality

$$a\left(1 - a^{\frac{1}{48}}\right) \geqslant .000138 .$$

We find that this inequality is satisfied for

$$.00104 \leqslant a \leqslant .99335 .$$

$$\text{Thus, Prob. } {}_{pp}(\xi < .000138) = \frac{50(49)}{2}(.000138)\left[\text{Log}_e\left(\frac{.99335}{.00104}\right) + 48(.866667 - .99986)\right]$$

$$= .16905\,(6.86188 - 6.39307) = .0793.$$

$$\therefore \text{ Prob. } {}_{pp}(R_n < .0044) = .0793.$$

The complementary part of the probability is given by

$$\int_0^{a_1} f(y)dy + \int_{a_2}^{1} f(y)dy = 1 + F(a_1) - F(a_2) \text{ , where}$$

$$F(y) = \int_0^y \frac{n(n-1)}{2}\left(1 - y^{\frac{1}{n-2}}\right)^2 dy = \frac{n(n-1)}{2}\,y - n(n-2)y^{\frac{n-1}{n-2}} + \frac{(n-1)(n-2)}{2}\,y^{\frac{n}{n-2}}.$$

Hence, the complementary part of the probability is

$$\text{Prob. } {}_{cp}(R_n < .0044) = 1 + F(.001040) - F(.993350)$$

$$= 1 + 1225(.001040) - 2400(.001040)^{\frac{49}{48}} + 1176(.001040)^{\frac{50}{48}} - 1225(.993350) + 2400(.993350)^{\frac{49}{48}} - 1176(.993350)^{\frac{50}{48}}$$

$$= .0291 .$$

$$\therefore \text{ Prob. } (R_n < .0044) = .0793 + .0291 = .1084.$$

Thus, if we decide to reject the hypothesis that the sample in Figure 1 comes from the population represented by line AB because the value of R_n is as small as .0044, the probability that our decision is wrong is .1084.

It can be seen that although the probability criterion R_n is the smallest of the three criteria thus far considered, it still is not as good a criterion as Pn or Qn, because, as we noted before,

$$\text{Prob. } (P_n < .0127) = .0925$$

and,

$$\text{Prob. } (Q_n < .0386) = .0386.$$

For this particular sample, therefore, the criterion Q_n is the best of the three thus far considered, since it yields the smallest probability of a wrong decision.

Miscellaneous Other Criteria and Their Distribution Functions

I. The Criterion $(h_1 - h_2)$

Let r_1 = rank of first failure
r_2 = rank of second failure
r_3 = rank of third failure

Furthermore, let h_1 = the difference be rank between second and first failures

$$= r_2 - r_1$$

and let h_2 = the difference in rank between third and second failures

$$= r_3 - r_2$$

Now take the statistic $\delta = h_1 - h_2$ as a criterion. In order to determine the distribution function of δ we must know the probability density function of h_2 and also the probability that $h_1 < \lambda$ if $a \leqslant h_2 \leqslant a + da$. The probability density function of $h_2 = r_3 - r_2$ is found as follows:

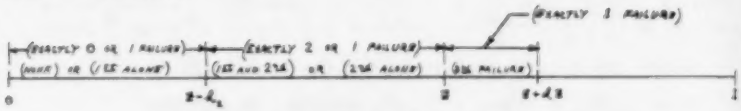


Figure 5

Prob. (Third failure lies between ranks z and z + dz) = $g_3(z)dz$, where $g_3(z)$ is the probability density function of the rank of the third failure.

Prob. (Second failure lies between $z-h_2$ and z if the third failure is between z and z + dz)

$$= \frac{2h_2(z-h_2)}{z^2} + \frac{h_2^2}{z^2} = \frac{h_2(2z-h_2)}{z^2} \qquad (\text{For } z > h_2)$$

$$= 1 \qquad (\text{For } z \leqslant h_2) \ .$$

Hence, Prob. (Second failure lies between $z-h_2$ and z, and that the third failure lies between z and z + dz)

$$= \frac{h_2(2z-h_2)}{z^2}\, g_3(z)dz \qquad (\text{For } z > h_2)$$

$$= g_3(z)dz \qquad (\text{For } z \leqslant h_2) \ .$$

$\therefore$ Prob. (Second failure is not more than h_2 rank units below the third failure)

$$= \int_0^{h_2} g_3(z)dz + \int_{h_2}^1 \frac{h_2(2z-h_2)}{z^2}\, g_3(z)dz .$$

Now $$g_3(z) = \frac{n(n-1)(n-2)}{2!}\, z^2\,(1-z)^{n-3} .$$

$$\therefore \text{Prob.}\,(r_2 - r_2 < h_2) = \int_0^{h_2} \frac{n(n-1)(n-2)}{2}\, z^2\,(1-z)^{n-3}\, dz$$

$$+ \int_{h_2}^1 \frac{h_2 n(n-1)(n-2)}{2}\,(2z - h_2)(1-z)^{n-3}\, dz$$

$$= \left[1 - (1-z)^n - nz(1-z)^{n-1} - \frac{n(n-1)}{2}\, z^2\,(1-z^{n-2}\right]_0^{h_2}$$

$$+ \frac{h_2 n(n-1)(n-2)}{2} \left[- \frac{(2z-h_2)(1-z)^{n-2}}{n-2} - \frac{2}{(n-1)(n-2)}\,(1-z)^{n-1}\right]_{h_2}^{1}$$

$$= 1 - (1-h_2)^n - nh_2(1-h_2)^{n-1} - \frac{n(n-1)}{2}\, h_2^2(1-h_2)^{n-2}$$

$$+ \frac{n(n-1)}{2}\, h_2^2(1-h_2)^{n-2} + nh_2(1-h_2)^{n-1}$$

$$= 1 - (1-h_2)^n .$$

Therefore, the probability density function of h_2 is

$$f(h_2) = n(1-h_2)^{n-1} .$$

In order to determine the probability that $h_1 < \mu$ if $a \leq h_2 \leq a + da$ we first prove the following:

LEMMA: The probability that the difference in rank between the third failure and the first failure is not more than k is $1 - (1-k)^n - n\,k\,(1-k)^{n-1}$.

Proof

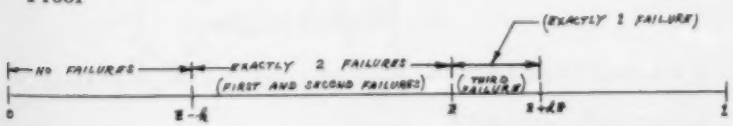


Figure 6

Prob. (Third failure lies between ranks z and z + dz) = $g_3(z)dz$ where $g_3(z)$ is the probability density function of the rank of the third failure.

Prob. [(Exactly 2 failures between z-k and z) if (third failure lies between z and z + dz)]

$$= \frac{k^2}{z^2} \quad \text{if} \quad z > k$$

$$= 1 \quad \text{if} \quad z \leqslant k \ .$$

Prob. [(Third failure lies between z and z + dz) and (exactly 2 failures between z-k and z)]

$$= \frac{k^2 g_3(z)}{z^2}\, dz \quad \text{if} \quad z > k$$

$$= g_3(z)dz \quad \text{if} \quad z \leqslant k \ .$$

$\therefore$ Prob. (Difference in rank between third and first failures is not more than k)

$$= \int_0^k g_3(z)dz + \int_k^1 \frac{k^2}{z^2}\, g_3(z)dz \ .$$

Since $g_3(z) = \frac{n(n-1)(n-2)}{2} z^2(1-z)^{n-3}$, this becomes

$$\text{Prob. } (r_3 - r_1 < k) = \int_0^k \frac{n(n-1)(n-2)}{2} z^2(1-z)^{n-3}\, dz$$

$$+ \int_k^1 \frac{k^2 n(n-1)(n-2)}{2} (1-z)^{n-3}\, dz$$

$$= \left[1 - (1-z)^n - nz(1-z)^{n-1} - \frac{n(n-1)}{2} z^2 (1-z)^{n-2}\right]_0^k$$

$$- \left[\frac{k^2 n(n-1)}{2} (1-z)^{n-2}\right]_k^1$$

$$= 1 - (1-k)^n - nk(1-k)^{n-1} - \frac{n(n-1)}{2} k^2 (1-k)^{n-2} + \frac{n(n-1)}{2} k^2 (1-k)^{n-2}$$

$$= 1 - (1-k)^n - nk(1-k^{n-1}$$

Q. E. D.

Having proved the lemma, we can now put $k = h_1 + h_2$.

Hence, Prob. $(r_3 - r_1 < h_1 + h_2)$

$$= 1 - (1-h_1 - h_2)^n - n(h_1 + h_2)(1-h_1 - h_2)^{n-1}$$

or, Prob. $\left[(r_2 - r_1 < h_1) \text{ if } (r_3 - r_2 = h_2)\right]$

$$= 1 - (1-h_1 - h_2)^n - n(h_1 + h_2)(1-h_1 - h_2)^{n-1}.$$

$\therefore$ Prob. $\left[(h_1 < \mu) \text{ if } (a \leqslant h_2 \leqslant a + da)\right]$

$$= 1 - (1-\mu - a)^n - n(\mu+a)(1-\mu-a)^{n-1} .$$

We have, therefore,

Prob. $(a \leqslant h_2 \leqslant a + da) = n(1-a)^{n-1}\, da$

Prob. $\left[(h_1 < a + \delta_0) \subset (a \leqslant h_2 \leqslant a + da)\right]$

$$= 1 - (1-\delta_0 - 2a)^n - n(\delta_0 + 2a)\,(1-\delta_0 - 2a)^{n-1}$$

$\therefore$ By multiplication,

Prob. $\left[(h_1 < a + \delta_0) \cdot (a \leqslant h_2 \leqslant a + da)\right]$

$$= n(1-a)^{n-1}\left[1-(1-\delta_0 - 2a)^n - n(\delta_0 + 2a)\,(1-\delta_0 - 2a)^{n-1}\right] da$$

or Prob. $\left[(h_1 - h_2 < \delta_0) \cdot (a \leqslant h_2 \leqslant a + da)\right]$

$$= n(1-a)^{n-1}\left[1 - (1-\delta_0 - 2a)^n - n(\delta_0 + 2a)(1-\delta_0 - 2a)^{n-1}\right] da$$

$\therefore$ Prob. $_{pp}(h_1 - h_2 < \delta_0)$

$$= \int_{a_1}^{a_2} n(1-a)^{n-1}\left[1 - (1-\delta_0 - 2a)^n - n(\delta_0 + 2a)(1-\delta_0 - 2a)^{n-1}\right] da$$

$$(1) \qquad = (1-a_1)^n - (1-a_2)^n + n(n-1)\int_{a_1}^{a_2} (1-a)^{n-1}(1-\delta_0 - 2a)^n\, da$$

$$- n^2 \int_{a_1}^{a_2} (1-a)^{n-1}(1-\delta_0 - 2a)^{n-1}\, da$$

where a_1 and a_2 are determined from the inequalities

$$o < 1 - (1-\delta_o - 2a)^n - n(\delta_o + 2a)(1-\delta_o - 2a)^{n-1} < 1 .$$

For $o < \delta_o < 1$ and n even, these inequalities are satisfied in the interval

$$o < a < \frac{1-\delta_o}{2}.$$

Hence, in case $o < \delta_o < 1$ and n is even, we have

$$\begin{cases} a_1 = o \\ a_2 = \dfrac{1-\delta_o}{2} . \end{cases}$$

For the sample given in Figure 1, we have

$$\begin{aligned} h_1 &= .074 - .011 = .063 \\ h_2 &= .087 - .074 = .013 \\ n &= 50 \end{aligned}$$

The criterion $h_1 - h_2$ becomes $.063 - .013 = .050$.

We now might ask, How significantly large is the value of $h_1 - h_2 = .050$? In other words, what are the chances that $h_1 - h_2 = .050$ is ever exceeded? To answer this question we go to formula (1), and obtain (taking $a_1 = 0$ and $a_2 = .475$).

$$\text{Prob.}_{pp}(h_1 - h_2 < .050) = 1 - (1-.475)^{50}$$

$$+ 2450\int_0^{.475} (1-a)^{49}(.95-2a)^{50}\,da - 2500\int_0^{.475} (1-a)^{49}(.95-2a)^{49}\,da .$$

The integrals in this expression are best evaluated by some accurate method of numerical integration, such as Weddle's rule. We find that

$$\int_0^{.475} (1-a)^{49}(.95 - 2a)^{49}\,da = .000525924$$

$$\int_0^{.475} (1-a)^{49}(.95 - 2a)^{50}\,da = .000492878$$

$$\therefore \text{Prob.}_{pp}(h_1 - h_2 < .050) = 1 - (.525)^{50} + 2450\,(.000492878) - 2500\,(.000525924)$$
$$= 1 - .1073$$
$$= .8927 .$$

In order to determine the complementary part of the probability, we have

$$F(y) = \int_0^y n(1-y)^{n-1}\,dy = 1 - (1-y)^n .$$

Hence, $\text{Prob.}_{cp}(h_1 - h_2 < .050) = 1 + F(o) - F(.475)$
$$= 1 + o - \left[1 - (.525)^{50}\right]$$
$$= .0000 +$$

$\therefore$ Prob. $(h_1 - h_2 < .050) = .8927$, or Prob. $(h_1 - h_2 > .050) = .1073$.

This last number tells us the significance of the value $h_1 - h_2 = .050$. We see that if we reject the hypothesis that the sample of Figure 1 comes from the population represented by line AB, the probability that we are wrong is .1073, since it is theoretically possible for 10.73 per cent of all samples of size 50 drawn from the population to have a value $h_1 - h_2 \geq .050$. Hence, for this particular sample, the criterion $\delta = h_1 - h_2$ is also inferior to the quantity Q_n of Example II.

II. The Criterion $\lambda = \dfrac{r_j}{r_{j+1}}$ $\quad (1 \leq j \leq n-1)$

Let r_j = rank of the j^{th} failure

r_{j+1} = rank of the $(j+1)^{th}$ failure,

then consider the statistic defined by the ratio $\lambda = \dfrac{r_j}{r_{j+1}}$. What is its distribution function? To answer this equation we must know the probability density function of r_{j+1}, and the probability that

$$r_j < \lambda_o \quad \text{if} \quad a \leq r_{j+1} \leq a + da .$$

It has been shown that the probability density function of r_{j+1} is

$$g(z) = \frac{n!}{j!\,(n-j-1)!}\, z^j\, (1-z)^{n-j-1},$$

where z represents r_{j+1} .

(See *Industrial Mathematics* Vol. 2, p. 3, Formula (A).)

Furthermore, it can be seen from Figure 7 that

$$\text{Prob.}\left[\begin{array}{l}\text{(Exactly j failures below rank } \lambda_o) \text{ if} \\ \text{(j + 1)}^{\text{th}} \text{ failure lies between z and z + dz)}\end{array}\right] = \frac{\lambda_o^{\,j}}{z^j}\ .$$

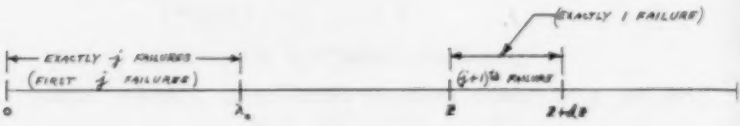


Figure 7

Thus $\text{Prob.}\ (a \leqslant r_{j+1} \leqslant a + da) = \frac{n!}{j!(n-j-1)!}\, a^j\, (1-a)^{n-j-1}\, da$

and $\text{Prob.}\left[(r_j < \lambda_o) \subset (a \leqslant r_{j+1} \leqslant a + da)\right] = \frac{\lambda_o^{\,j}}{a^j}$

$\therefore$ $\text{Prob.}\left[(r_j < a\lambda_o) \subset (a \leqslant r_{j+1} \leqslant a + da)\right] = \lambda_o^{\,j}$.

By multiplication: $\text{Prob.}\left[(r_j < a\lambda_o) \cdot (a \leqslant r_{j+1} \leqslant a + da)\right]$

$$= \frac{n!}{j!\,(n-j-1)!}\lambda_o^{\,j}\, a^j\, (1-a)^{n-j-1}\, da$$

or, $\text{Prob.}\left[\left(\frac{r_j}{r_{j+1}} < \lambda_o\right) \cdot (a \leqslant r_{j+1} \leqslant a + da\right]$

$$= \frac{n!}{j!\,(n-j-1)!}\lambda_o^{\,j}\, a^j\, (1-a)^{n-j-1}\, da$$

$$\therefore\ \text{Prob.}\left(\frac{r_j}{r_{j+1}} < \lambda_o\right) = \int_0^1 \frac{n!}{j!(n-j-1)!}\lambda_o^{\,j}\, a^j\, (1-a)^{n-j-1}\, da = \lambda_o^{\,j}\ .$$

Hence, the statistic $\lambda = \frac{r_j}{r_{j+1}}$ has a cumulative distribution function $F(\lambda) = \lambda^j$, i.e., a probability density function $f(\lambda) = j\lambda^{j-1}$.

This criterion λ turns out to be very useful in case two successive upper extreme values are widely separated. Let us say, for example, that in a sample of 51 specimens the 50th failure is at rank .920, and the 51st failure is at rank .980 of a given population. Then, taking $j = 50$,

$$\lambda = \frac{r_j}{r_{j+1}} = \frac{r_{50}}{r_{51}} = \frac{.920}{.980} = .9387755$$

and Prob. $(\lambda < .9387755) = (.9387755)^{50} = .0425$.

This indicates that such a sample would not readily come from the population in question, since a value of λ as low as .9387755 is to be expected only in about four samples out of a hundred.

6

Analy

Vehic

By J
Mech
Armo
Instit
Chic

Pr
dynam
niques
the re
theory
drawn
be use
to give
vehicl
can be
being
clusio
pensio
summ

As
speeds
vehicl
comfo
more
vehicl
transp
proble
which
requir

In
respon
litera
made

*Base
quiren
of Tec

6

Analytical and Experimental Studies of Vehicle Dynamic Behavior*

By John Arthur Carlson
Mechanical Engineering Research Department
Armour Research Foundation of Illinois
Institute of Technology
Chicago

SUMMARY

Progress in the development of analysis in the field of vehicle dynamics has been made possible by modern computational techniques. In this study, an appraisal of analysis is made by comparing the results of analysis with actual vehicle tests, both experiment and theory being included in one investigation. The general conclusion drawn from such an appraisal is that relatively simple systems can be used to represent a vehicle in motion over a variety of courses to give good quantitative results; and further, the effects of varying vehicle properties (spring rates, damping characteristics, masses) can be studied by using these simple systems, the results obtained being of sufficient accuracy to be used for design purposes. Conclusions reached on the effects of some automobile properties (suspension bump stops, suspension damping, and tire damping) are summarized.

INTRODUCTION

As the ability of ground vehicles to attain and sustain higher speeds has increased in recent years, the dynamic behaviour of such vehicles has become more critical, and the problems of passenger comfort and safety, and vehicle stability and control have become more complex and difficult. Basically, the problem is to design a vehicle to satisfy the requirements of the personnel or material transported. The goal is to obtain the optimum solution to this problem in the most economical way. Analytical design techniques which may be applied to the vehicle to predict its behaviour are required.

In the work described here, the problem of obtaining the dynamic response of a vehicle was studied with the above goal in mind. The literature on this subject reveals that many analyses have been made from which general conclusions regarding the dynamic

*Based on a thesis submitted in May, 1955, in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the California Institute of Technology.

behaviour of vehicles have been drawn.[1,2,3] Practically all of the analyses are limited to linear systems, and in only a very few cases is the response of a system computed. Most papers involve only a discussion of the qualitative effects of various terms which appear in the equation of motion, the discussion sometimes being based on only the one degree-of-freedom system concept. In practice, an attempt is made to apply the meager analytical conclusions available; however, most designs evolve through testing. With the advent of modern computational techniques, it is now possible to reappraise the problem of analyzing a vehicle and predicting its behaviour.[4,5]

VEHICLE TESTING AND COMPUTER DEVELOPMENTS

The development of modern instrumentation has made possible accurate and extensive determination of vehicle dynamic behaviour from testing.[6,7,8,9] While much information can be gained experimentally, it may be seen that purely experimental design studies in which vehicle properties are to be varied must be limited within economic bounds. These limited studies give an incomplete picture and sometimes lead to inconclusive results.

Another approach to the problem of vehicle design is to analyze the vehicle as a mechanical system of many degrees of freedom which represents the motions of the vehicle components, such as wheels, frame, body, engine, and passengers. The troubles which face a theoretical analysis can be seen by considering an example. A system which can be used to represent an automobile and its suspension approximately for small vertical motions is shown in Figure 1. Many simplifying assumptions were necessary to arrive at this system, such as the assumption that the engine and the body and its contents are rigidly connected to the frame which is rigid and that the tire contacts the ground at a point directly below the wheel. The equations of motion of this system are as follows:

$$m_1\ddot{x}_1 - G_1\left(\dot{x}_5 + \frac{d}{2}\dot{\theta} + a\dot{\phi} - \dot{x}_1\right) - K_1\left(x_5 + \frac{d}{2}\theta + a\phi - x_1\right) - \frac{K_8}{d}\left(\theta - \frac{x_1 - x_2}{d}\right)$$

$$+ K_3(x_1 - y_1) = 0; \tag{1}$$

$$m_2\ddot{x}_2 - G_2\left(\dot{x}_5 - \frac{d}{2}\dot{\theta} + a\dot{\phi} - \dot{x}_2\right) - K_2\left(x_5 - \frac{d}{2}\theta + a\phi - x_2\right)$$

$$+ \frac{K_8}{d}\left(\theta - \frac{x_1 - x_2}{d}\right) + K_4(x_2 - y_2) = 0; \tag{2}$$

$$\left(\frac{m_3}{4} + \frac{I_3}{e^2}\right)\ddot{x}_3 + \left(\frac{m_3}{4} - \frac{I_3}{e^2}\right)\ddot{x}_4 - G_3\left(\dot{x}_5 + \frac{e}{2}\dot{\theta} - b\dot{\phi} - \dot{x}_3\right) - K_3\left(x_5 + \frac{e}{2}\theta - b\phi - x_3\right)$$

$$(3)\qquad -\frac{K_{10}}{e}\left(\theta-\frac{x_3-x_4}{e}\right)+K_7(x_3-y_3)=0;$$

$$\left(\frac{m_3}{4}-\frac{I_3}{e^2}\right)\ddot{x}_3+\left(\frac{m_3}{4}+\frac{I_3}{e^2}\right)\ddot{x}_4-G_4\left(\dot{x}_5-\frac{e\dot{\theta}}{2}-b\dot{\phi}-\dot{x}_4\right)-K_4\left(x_5-\frac{e}{2}\theta-b\phi-x_4\right)$$

$$(4)\qquad +\frac{K_{10}}{e}\left(\theta-\frac{x_3-x_4}{e}\right)+K_8(x_4-y_4)=0;$$

$$m_4\ddot{x}_5+G_1\left(\dot{x}_5+\frac{d}{2}\dot{\theta}+a\dot{\phi}-\dot{x}_1\right)+G_2\left(\dot{x}_5-\frac{d}{2}\dot{\theta}+a\dot{\phi}-\dot{x}_2\right)+G_3\left(\dot{x}_5+\frac{e}{2}\dot{\theta}-b\dot{\phi}-\dot{x}_3\right)$$

$$+G_4\left(\dot{x}_5-\frac{e}{2}\dot{\theta}-b\dot{\phi}-\dot{x}_4\right)+K_1\left(x_5+\frac{d}{2}\theta+a\phi-x_1\right)$$

$$+K_2\left(x_5-\frac{d}{2}\theta+a\phi-x_2\right)$$

$$(5)\qquad +K_3\left(x_5+\frac{e}{2}\theta-b\phi-x_3\right)+K_4\left(x_5-\frac{e}{2}\theta-b\phi-x_4\right)=0;$$

$$I_1\ddot{\phi}+aG_1\left(\dot{x}_5+\frac{d}{2}\dot{\theta}+a\dot{\phi}-\dot{x}_1\right)+aG_2\left(\dot{x}_5-\frac{d}{2}\dot{\theta}+a\dot{\phi}-\dot{x}_2\right)$$

$$-bG_3\left(\dot{x}_5+\frac{e}{2}\dot{\theta}-b\dot{\phi}-\dot{x}_3\right)-bG_4\left(\dot{x}_5-\frac{e}{2}\dot{\theta}-b\dot{\phi}-\dot{x}_4\right)$$

$$+aK_1\left(x_5+\frac{d}{2}\theta+a\phi-x_1\right)+aK_2\left(x_5-\frac{d}{2}\theta+a\phi-x_2\right)$$

$$(6)\qquad -bK_3\left(x_5+\frac{e}{2}\theta-b\phi-x_3\right)-bK_4\left(x_5-\frac{e}{2}\theta-b\phi-x_4\right)=0;$$

$$I_2\ddot{\theta}+\frac{d}{2}G_1\left(\dot{x}_5+\frac{d}{2}\dot{\theta}+a\dot{\phi}-\dot{x}_1\right)-\frac{d}{2}G_2\left(\dot{x}_5-\frac{d}{2}\dot{\theta}+a\dot{\phi}-\dot{x}_2\right)$$

$$+\frac{e}{2}G_3\left(\dot{x}_5+\frac{e}{2}\dot{\theta}-b\dot{\phi}-\dot{x}_3\right)$$

$$-\frac{e}{2}G_4\left(\dot{x}_5-\frac{e}{2}\dot{\theta}-b\dot{\phi}-\dot{x}_4\right)+\frac{d}{2}K_1\left(x_5+\frac{d}{2}\theta+a\phi-x_1\right)$$

$$-\frac{d}{2}K_2\left(x_5-\frac{d}{2}\theta+a\phi-x_2\right)$$

$$+\frac{e}{2}K_3\left(x_5+\frac{e}{2}\theta-b\phi-x_3\right)-\frac{e}{2}K_4\left(x_5-\frac{e}{2}\theta-b\phi-x_4\right)$$

$$(7)\qquad +K_9\left(\theta-\frac{x_1-x_2}{d}\right)+K_{10}\left(\theta-\frac{x_3-x_4}{e}\right)=0.$$

Obtaining solutions to these equations by conventional analytical methods would not be feasible unless many more simplifying assumptions were made.[10] As simplifying assumptions are made the solutions may become increasingly unrealistic and the information obtained becomes more uncertain. The next step is to turn to the use of an electrical computer; but this is not as simple as it sounds. Solving the above equations, which by no means represent the vehicle completely, with the addition of complications such as bump stops and non-linear shock absorbers, would tax the capacity of even the largest computers available. Even if solutions to these equations were obtained, one still would not be certain of how well the actual vehicle would have been represented. These computers are very costly and require specialized personnel to operate them. That both of these factors tend to hinder their ready use by engineers is a point which may be overlooked by computer designers. Besides

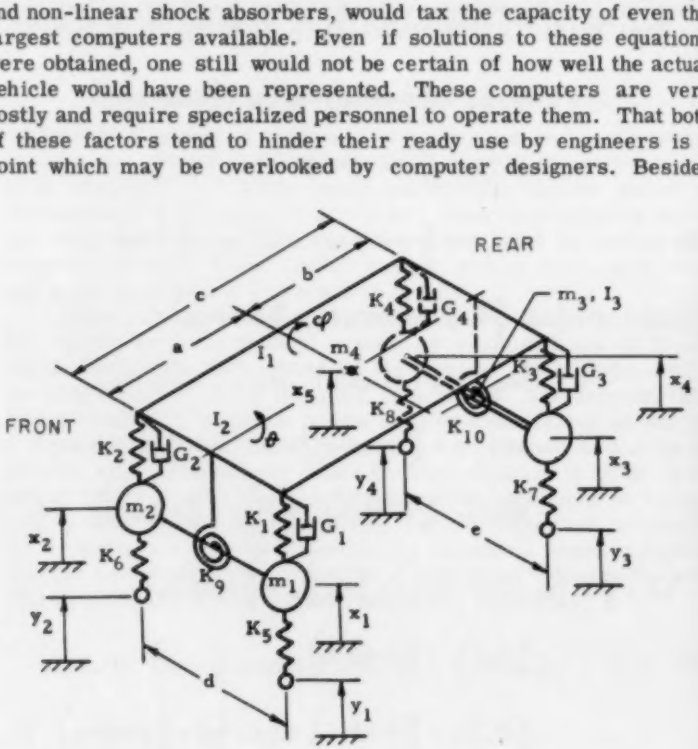


Figure 1

Seven Degree-Of-Freedom System for Representing Small Vertical Motions of an Automobile: m_4 - Chassis and Body; m_1, m_2 - Front Wheels; m_3 - Rear Axle; K_1, K_2 - Front Suspension Springs; K_3, K_4 - Rear Suspension Springs; K_5, K_6, K_7, K_8, - Tires; G - Shock Absorbers; K_9, K_{10} - Roll Stabilizers.

this, the given data may not be known with sufficient accuracy, or extreme accuracy may not be necessary for design purposes, to justify an elaborate computer setup.

In an effort to resolve the above difficulties in testing and analysis, a research program was set up, the main object of which was to conduct an appraisal of analysis by comparing the results of analysis and actual vehicle tests. It was hoped that simple mechanical systems (simplified over that represented in Figure 1) could be used to represent the vehicle and compute its response to a given excitation. A previous work along this line[11] was very limited in scope, and the analysis was performed independently of the vehicle tests. The work described here is the first known to the author which included performing the vehicle tests as well as the analysis. Including both experiment and theory in one investigation allowed a complete study of all methods and procedures used. This is a necessity if the evaluation of the analysis and vehicle tests is to be valid.

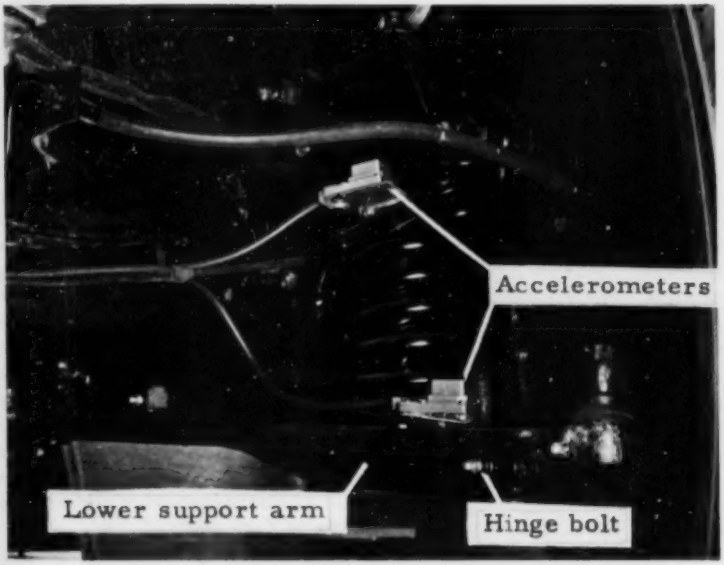


Figure 2

Installation of Accelerometers at the Right Front Wheel.

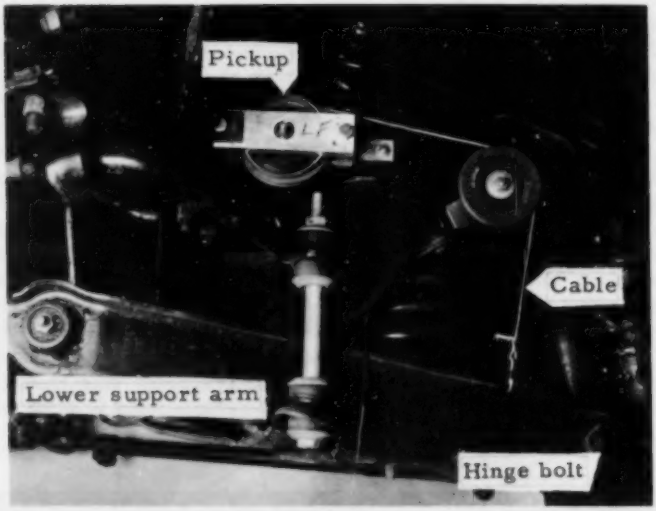


Figure 3

Installation of Displacement Pickup at the Left Front Wheel.

VEHICLE TESTS

A standard production automobile was instrumented so that the vertical accelerations of the frame and axle at each wheel, and the relative vertical displacement between each wheel and the frame could be measured. Using acceleration as a measure assured an adequate description of the response, since velocities and displacements can be obtained accurately from the acceleration by integration. Further, the significance of any higher derivatives in vehicle rideability is not well established.[12] The installation of accelerometers at the right front wheel is shown in Figure 2. Figure 3 shows the displacement pickup installed at the left front wheel. Accelerometers and the displacement pickup installed at the right rear wheel are shown in Figures 4 and 5. Speed and distance travelled were measured by means of a fifth wheel. The arrangement or recording equipment within the test car is shown in Figure 6.*

Courses were laid out over which the vehicle was driven in a straight line at constant speeds. One of these courses contained a large dip which excited low frequency symmetric motion of the body (bounce and pitch motions). A profile of this course is shown in

*For a complete description of the instrumentation, see the thesis referred to in the note at the beginning of this article.

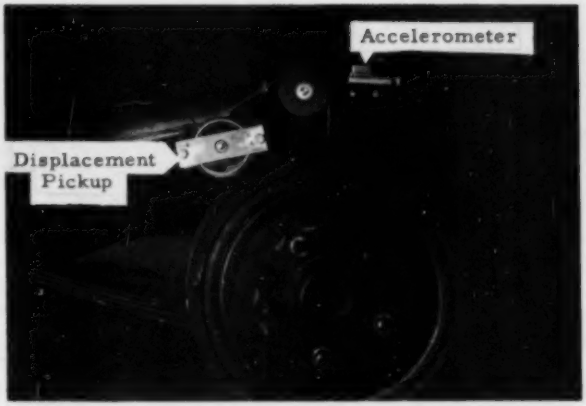


Figure 4

Displacement Pickup and Accelerometer Mounted on the Frame at the Right Rear Wheel.

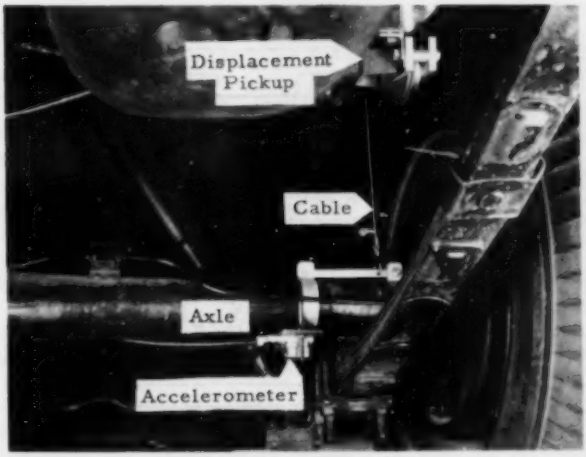


Figure 5

Displacement Pickup and Axle Mounted Accelerometer at the Right Rear Wheel.

Figure 6

Arrangement of Recording Equipment in the Test Car.

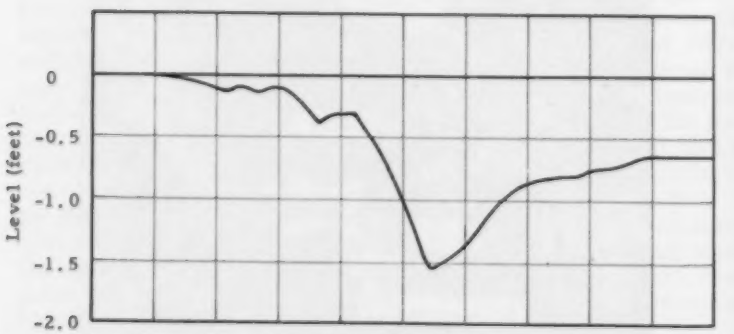


Figure 7

Profile of Test Dip.

Figure 7. Higher frequency motions of the wheels (wheel bounce) were excited by driving the car over short bumps of wood which were placed on a smooth pavement. These are shown in Figure 8 in position to excite right and left wheels at the same time (symmetric motion). By removing one of the bumps the wheels on only one side of the car could be excited.

Records of the accelerations of the frame and wheels, and of the relative displacements of the wheels with respect to the frame, were made while driving the test car over these courses. Figure 9 shows a typical record made in traveling over the test dip. Typical vehicle test records obtained while going over one of the wooden bumps are shown in Figure 10.

A four degree-of-freedom system which could be used to represent the motion of the vehicle in symmetric motion is shown in Figure 11, together with the differential equations of motion which describe it.

For the vehicle used, the front to rear mass coupling term M_{12} (see Figure 11) was found to be small compared with the other mass terms M_{11} and M_{22}, which is typical of most modern American cars. If this term is neglected in the analysis the system shown in Figure 11 reduces to two two degree-of-freedom systems. This two degree-of-freedom system, shown in Figure 12 together with its differential equations of motion, was used to obtain the analytical results for bounce and pitch motions.

An examination of Figure 10 reveals that wheel bounce of each of the front wheels is practically independent while the motion of left and right rear wheels is strongly coupled. A system which was used to compute the motion of the rear axle in higher frequency asymmetric motion is shown in Figure 13 with the corresponding differential equations of motion.* It is assumed that this motion of the rear axle is independent of the higher frequency motion of the front wheels.

To obtain solutions to the differential equations of motion of the above systems for arbitrary input functions in a reasonable time, computers are a necessity, especially if the equations are nonlinear. The computed results presented in the next section were obtained by using a small electric analog computer, the principles of operation of which are presented in the literature.[13,14] Operation of this type of computer is based on the fact that the mathematical equations describing the reaction of electrical circuits containing inductors, resistors, and condensers, are the same as those for analogous mechanical systems. This type of computer is very suitable for studying the dynamic behaviour of mechanical systems. Other kinds of computers which can be used are the electronic differential analyzer (sometimes called "electronic analog computer"),[15] the mechanical differential analyzer,[16] and the digital computer.[17]

*For derivations of the equations in this section, see the thesis.

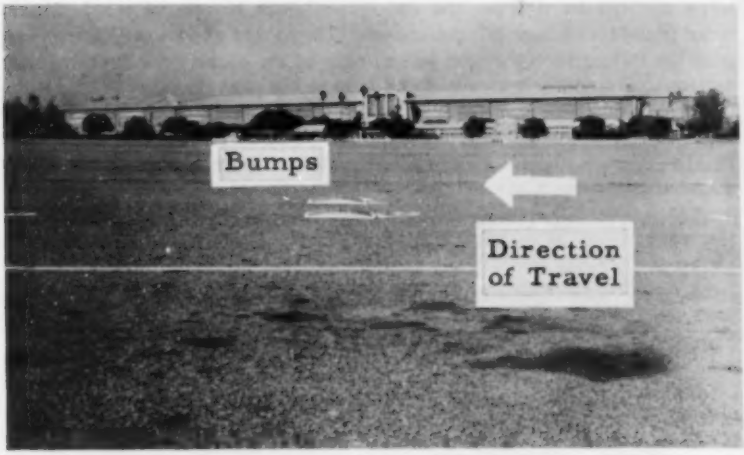


Figure 8

Side View of Wooden Bumps Placed Symmetrically.

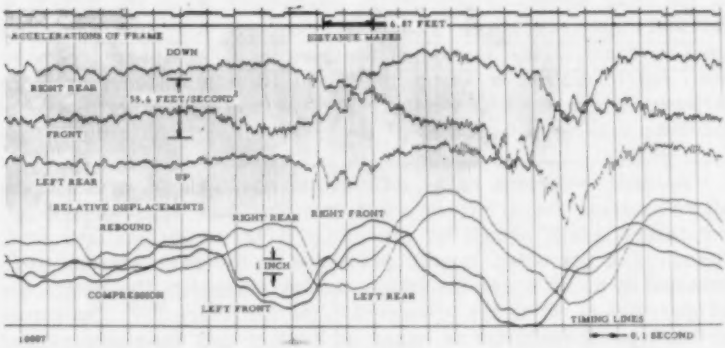
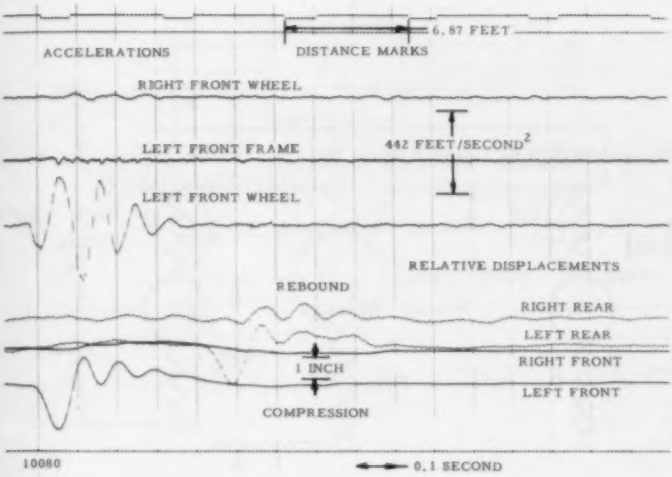
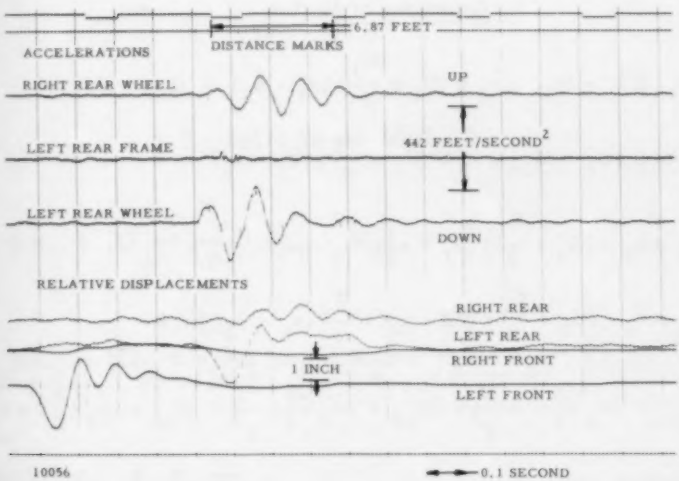


Figure 9

Typical Vehicle Test Record Obtained While Travelling Over the Test Dip at 25 mph.



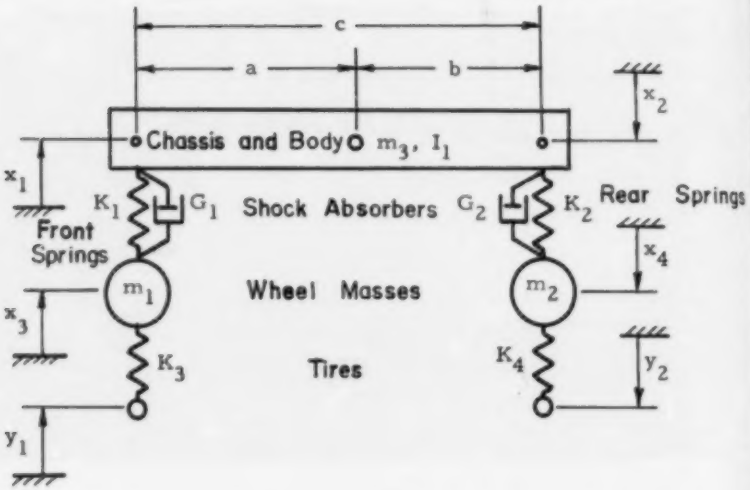
a



b

Figure 10

Vehicle Records Obtained in Going Over a Wooden Half-Sine Bump With the Left Wheels Only, at 15 mph.



Equations of Motion

$$M_{11}\, p\dot{x}_1 - M_{12}\, p\dot{x}_2 + (G_1 + \frac{K_1}{p})(\dot{x}_1 - \dot{x}_3) = 0$$

$$- M_{12}\, p\dot{x}_1 + M_{22}\, p\dot{x}_2 + (G_2 + \frac{K_2}{p})(\dot{x}_2 - \dot{x}_4) = 0$$

$$m_1\, p\dot{x}_3 - (G_1 + \frac{K_1}{p})(\dot{x}_1 - \dot{x}_3) + \frac{K_3}{p}(\dot{x}_3 - \dot{y}_1) = 0$$

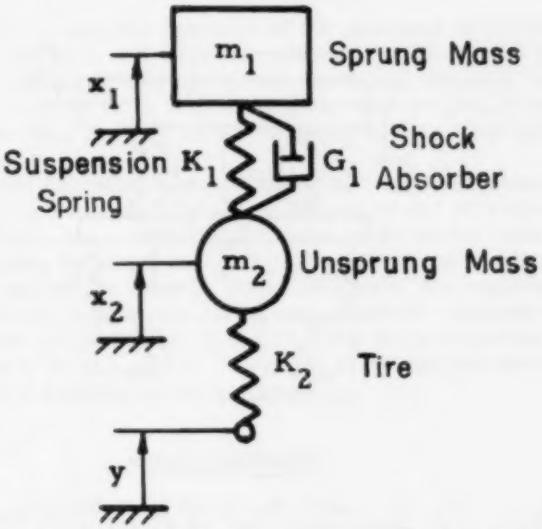
$$m_2\, p\dot{x}_4 - (G_2 + \frac{K_2}{p})(\dot{x}_2 - \dot{x}_4) + \frac{K_4}{p}(\dot{x}_4 - \dot{y}_2) = 0$$

$$M_{11} = m_3 (\frac{b}{c})^2 + \frac{I_1}{c^2}; \quad M_{12} = m_3 \frac{ab}{c^2} - \frac{I_1}{c^2}; \quad M_{22} = m_3 (\frac{a}{c})^2 + \frac{I_1}{c^2}$$

Operator, $p = \frac{d}{dt}$.

Figure 11

Four Degree-Of-Freedom System for Representing an Automobile in Symmetric Motion.



Equations of Motion

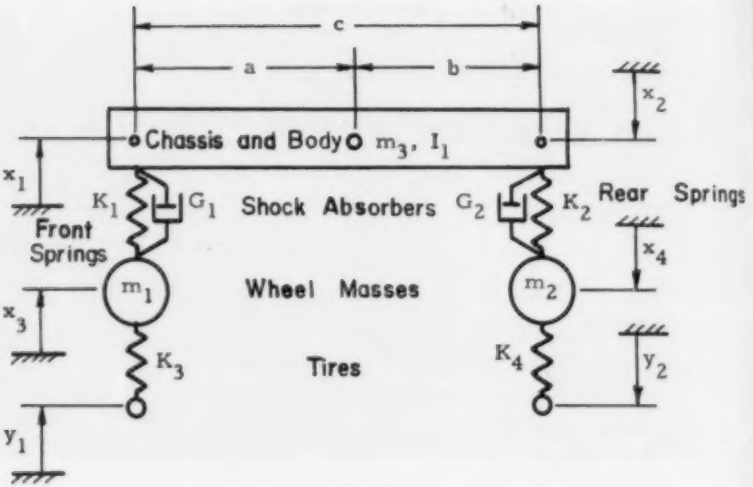
$$m_1 p\dot{x}_1 + (G_1 + \frac{K_1}{p})(\dot{x}_1 - \dot{x}_2) = 0$$

$$m_2 p\dot{x}_2 - (G_1 + \frac{K_1}{p})(\dot{x}_1 - \dot{x}_2) + \frac{K_2}{p}(\dot{x}_2 - \dot{y}) = 0$$

perator, $p = \frac{d}{dt}$.

Figure 12

Two Degree-Of-Freedom System for Representing an Automobile in Symmetric Motion.



Equations of Motion

$$M_{11}\, p\dot{x}_1 - M_{12}\, p\dot{x}_2 + (G_1 + \frac{K_1}{p})(\dot{x}_1 - \dot{x}_3) = 0$$

$$- M_{12}\, p\dot{x}_1 + M_{22}\, p\dot{x}_2 + (G_2 + \frac{K_2}{p})(\dot{x}_2 - \dot{x}_4) = 0$$

$$m_1\, p\dot{x}_3 - (G_1 + \frac{K_1}{p})(\dot{x}_1 - \dot{x}_3) + \frac{K_3}{p}(\dot{x}_3 - \dot{y}_1) = 0$$

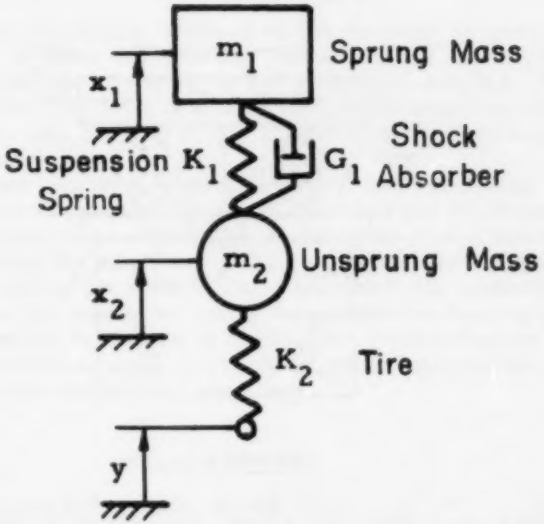
$$m_2\, p\dot{x}_4 - (G_2 + \frac{K_2}{p})(\dot{x}_2 - \dot{x}_4) + \frac{K_4}{p}(\dot{x}_4 - \dot{y}_2) = 0$$

$$M_{11} = m_3 (\frac{b}{c})^2 + \frac{I_1}{c^2}\,; \quad M_{12} = m_3 \frac{ab}{c^2} - \frac{I_1}{c^2}\,; \quad M_{22} = m_3 (\frac{a}{c})^2 + \frac{I_1}{c^2}$$

Operator, $p = \frac{d}{dt}$.

Figure 11

Four Degree-Of-Freedom System for Representing an Automobile in Symmetric Motion.



Equations of Motion

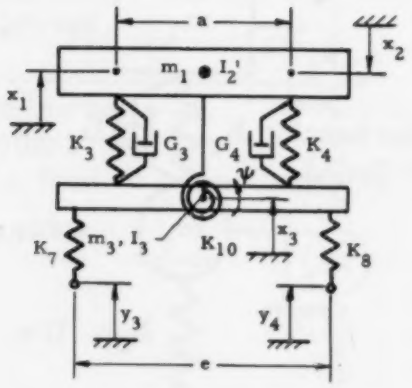
$$m_1 p\dot{x}_1 + (G_1 + \frac{K_1}{p})(\dot{x}_1 - \dot{x}_2) = 0$$

$$m_2 p\dot{x}_2 - (G_1 + \frac{K_1}{p})(\dot{x}_1 - \dot{x}_2) + \frac{K_2}{p}(\dot{x}_2 - \dot{y}) = 0$$

$$\text{Operator, } p = \frac{d}{dt} .$$

Figure 12

Two Degree-Of-Freedom System for Representing an Automobile in Symmetric Motion.



Equations of Motion

$$(\frac{m_1}{4}+\frac{I_2'}{a^2})p\,\dot{x}_1-(\frac{m_1}{4}-\frac{I_2'}{a^2})p\,\dot{x}_2+(G_3+\frac{K_3}{p})(\dot{x}_1-\dot{x}_3+\frac{a}{2}\dot{\psi})+\frac{K_{10}}{a^2p}(\dot{x}_1+\dot{x}_2+a\dot{\psi})=0.$$

$$-(\frac{m_1}{4}-\frac{I_2'}{a^2})p\,\dot{x}_1+(\frac{m_1}{4}+\frac{I_2'}{a^2})p\,\dot{x}_2+(G_4+\frac{K_4}{p})(\dot{x}_2+\dot{x}_3+\frac{a}{2}\dot{\psi})+\frac{K_{10}}{a^2p}(\dot{x}_1+\dot{x}_2+a\dot{\psi})=0.$$

$$m_3p\,\dot{x}_3-(G_3+\frac{K_3}{p})(\dot{x}_1-\dot{x}_3+\frac{a}{2}\dot{\psi})+(G_4+\frac{K_4}{p})(\dot{x}_2+\dot{x}_3+\frac{a}{2}\dot{\psi})+\frac{K_7}{p}(\dot{x}_3-\frac{e}{2}\dot{\psi}-\dot{y}_3)$$
$$+\frac{K_8}{p}(\dot{x}_3+\frac{e}{2}\dot{\psi}-\dot{y}_4)=0.$$

$$I_3p+\frac{a}{2}(G_3+\frac{K_3}{p})(\dot{x}_1-\dot{x}_3+\frac{a}{2}\dot{\psi})+\frac{a}{2}(G_4+\frac{K_4}{p})(\dot{x}_2+\dot{x}_3+\frac{a}{2}\dot{\psi})+\frac{K_{10}}{ap}(\dot{x}_1+\dot{x}_2+a\dot{\psi})$$
$$-\frac{eK_7}{2p}(\dot{x}_3-\frac{e}{2}\dot{\psi}-\dot{y}_3)+\frac{eK_8}{2p}(\dot{x}_3+\frac{e}{2}\dot{\psi}-\dot{y}_4)=0.$$

Operator, $p=\frac{d}{dt}$.

Figure 13

Four Degree-Of-Freedom System for Representing Motion of the Rear Axle: m_1 - Chassis and Body; m_3 - Rear Axle; K_3, K_4 - Rear Suspension Springs; K_7, K_8 - Tires; G_3, G_4 - Shock Absorbers; K_{10} - Roll Stabilizer.

VEHICLE PROPERTIES

In order to compute solutions of the equations of motion it is necessary to know the physical characteristics of the vehicle, masses, moments of inertia, spring rates, and damping. For the vehicle used in the tests described here, these properties were obtained from data supplied by the manufacturer or from laboratory measurements.

The front and rear suspension spring force-deflection curves were measured and are as shown in Figures 14 and 15, respectively. In the analysis, these were approximated by the dashed lines shown. Approximating the shock absorber force-velocity curves by straight lines and adding the average static friction of the suspension obtained from the suspension spring data, yields the damping characteristics shown in Figures 16 and 17. The force-deflection curves for the tire were as shown in Figure 18. In the analysis these curves were approximated by the dashed lines.

RESULTS

Low Frequency Motion of the Sprung Mass

The results of the vehicle tests and of the analysis are presented in the form of graphical comparisons of the responses of the vehicle obtained by the two methods.

Linear systems were used first in the computer studies in order to determine how much could be learned under this very radical simplifying assumption and also to provide a basis for comparison with non-linear systems subsequently studied. The assumption that the front-to-rear mass coupling term could be neglected when considering symmetric motion was checked on the computer and found to be valid.

Records showing the response of the vehicle in going over the test dip are reproduced in Figures 19 to 22. The measurements used in these comparisons are the relative displacements between the frame and the wheels, and the accelerations of the sprung mass m_1. Low frequency motion of the sprung mass on its suspension predominated, and higher frequency motion of the unsprung mass as well as noise present in the frame, largely due to the engine, were neglected. The vehicle measurements plotted are an average of those made on the left and right hand sides, since the motion was not exactly symmetric. The computer solutions appear only where they depart from the vehicle measurements. In these records the only non-linearity taken into account was that of the change in spring rate occurring when the main suspension spring bottomed on the rubber bump stops. Tire pressure was 24 psi gage, and the

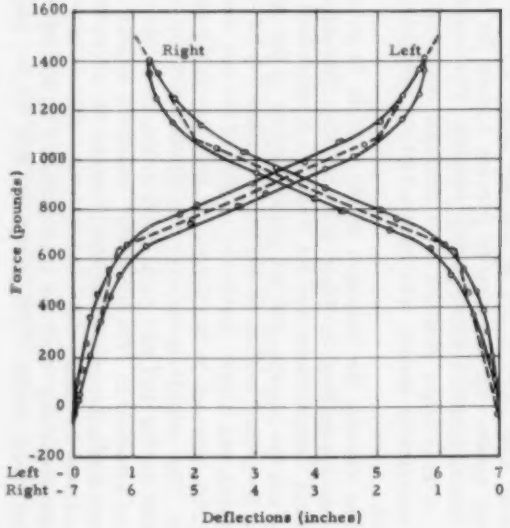


Figure 14

Front Suspension Spring Force-Deflection Curves.

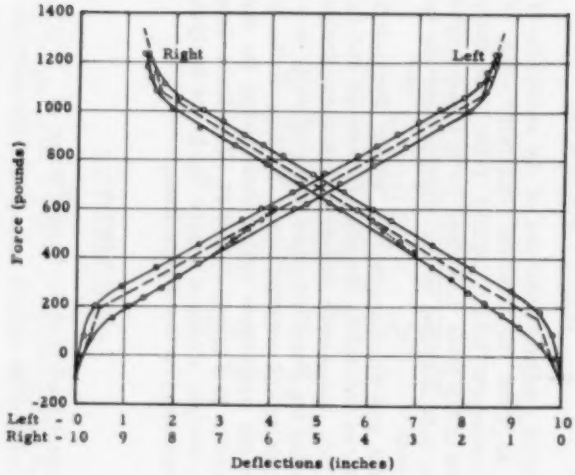


Figure 15

Rear Suspension Spring Force-Deflection Curves.

Front Suspension Spring Force-Deflection Curves.

Rear Suspension Spring Force-Deflection Curves.

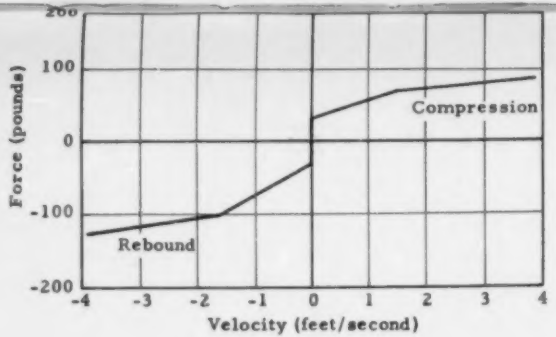


Figure 16

Front Suspension Damping Force-Velocity Curve.

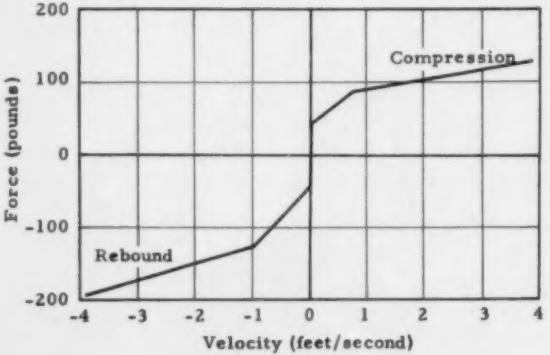


Figure 17

Rear Suspension Damping Force-Velocity Curve

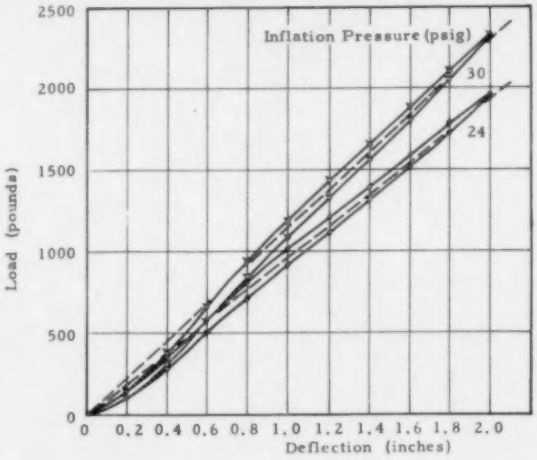


Figure 18

Tire Load-Deflection Curves.

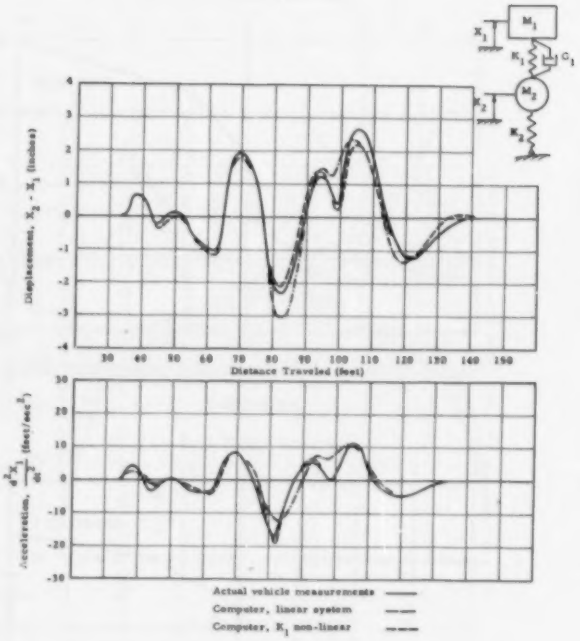


Figure 19

Motion at Front Wheels in Going Over Test Dip at 20 Miles Per Hour

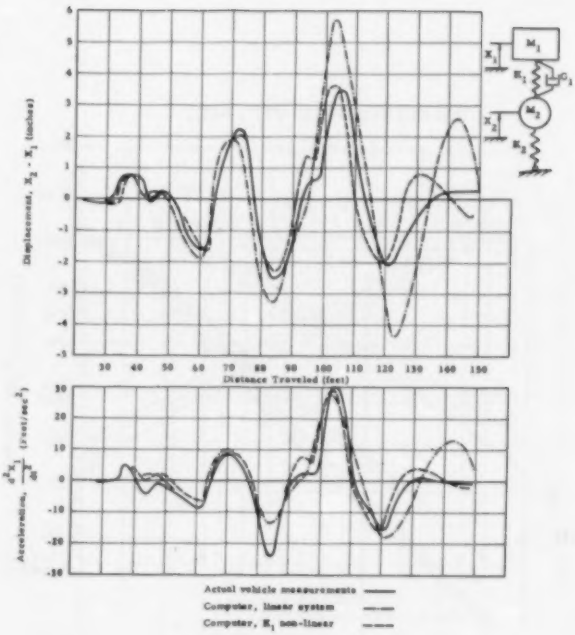
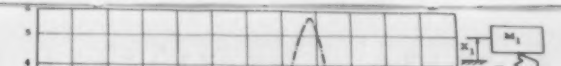


Figure 20

Motion at Front Wheels in Going Over Test Dip at 25 Miles Per Hour.



Motion at Front Wheels in Going Over Test Dip at 20 Miles Per Hour

Motion at Front Wheels in Going Over Test Dip at 25 Miles Per Hour.

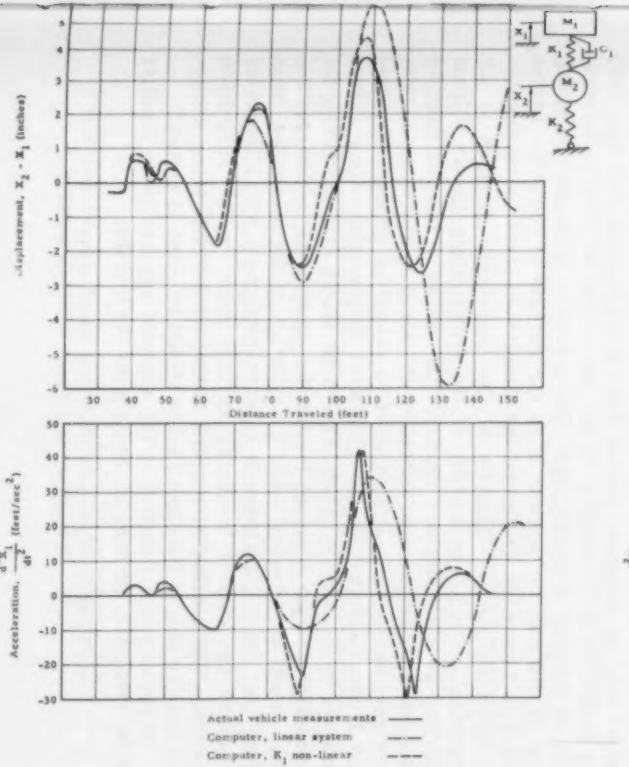


Figure 21

Motion at Front Wheels in Going Over Test Dip at 28 Miles Per Hour.

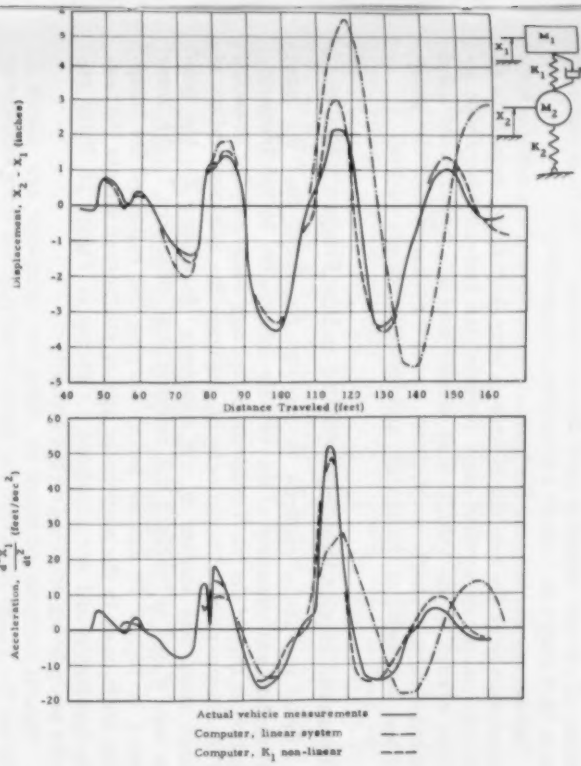


Figure 22

Motion at Rear Wheels in Going Over Test Dip at 25 Miles Per Hour.

values of viscous damping used for each wheel in the front and rear were 44.0 slug/sec. and 47.8 slug/sec., respectively.

The agreement between the computed results and the vehicle tests for motion over the test dip is surprisingly good if one considers all the sources of error that exist between the two curves. These include errors in the vehicle test measurements, measurements and calculation of vehicle properties, approximation of vehicle properties, surveying of the course, and overall computer error.

Effects of Changing Vehicle Properties on Low Frequency Motion

From the above records it may be concluded that the most influential non-linearity of the vehicle suspension in low frequency motion is that of the main suspension springs caused by the presence of the bump-stops. The effects of changing the characteristics of these stops was studied on the computer. In Figure 23, the pronounced effect of changing the rebound stop clearance on the lower peaks in Figure 19 can be seen. Varying the ratios of the bump-stop stiffnesses to the suspension spring stiffnesses affects the peak displacements and accelerations as shown in Figure 24, these results again being obtained on the computer.

Typical records obtained by using the computer are shown in Figure 25. These illustrate the effect of varying the main suspension spring rate on response motion when going over the test dip. Both the amplitude and frequency of the motion can be seen to increase with an increase in spring rate. The curves for a suspension spring rate of 106.2 lb/in. correspond to the non-linear computer curves shown in Figure 20. (Note that the direction of travel is reversed.)

Evidently the non-linearity of the suspension damping is not important in primary motion of the frame since linear viscous damping gave satisfactory results. The values of viscous damping used were chosen by approximating the damping characteristics by straight lines in the range of relative velocities between the wheel and frame indicated by the slopes of displacement-time curves obtained in the vehicle tests. Figure 26 shows that increasing the amount of viscous damping lowers the amplitudes of the motion. This would be expected since examination of the excitation when going over the test dip reveals a large component with a frequency close to the natural frequency sprung mass on the suspension springs. The conclusion reached is that for low frequency excitation the damping of the sprung mass should be close to critical.

Neglecting tire damping in the analysis appears to be justified by the results in the case of motion over the test dip.

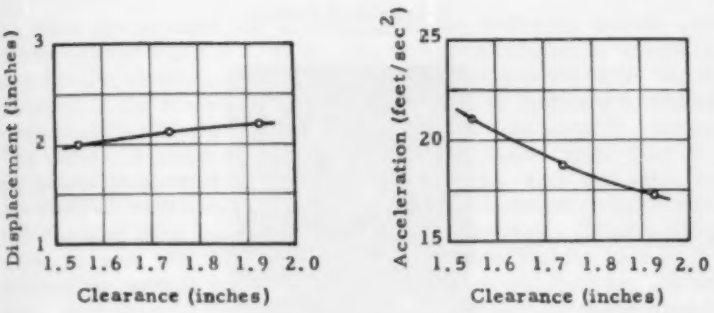


Figure 23

Effect of Varying the Rebound Stop Clearance at the Front Wheels in Going Over the Test Dip at 20 Miles Per Hour.

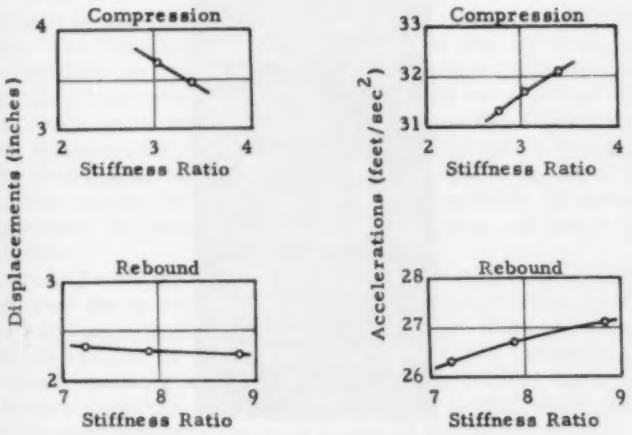


Figure 24

Effect of Varying Bump Stop Stiffnesses at the Front Wheels in Going Over the Test Dip at 25 Miles Per Hour.

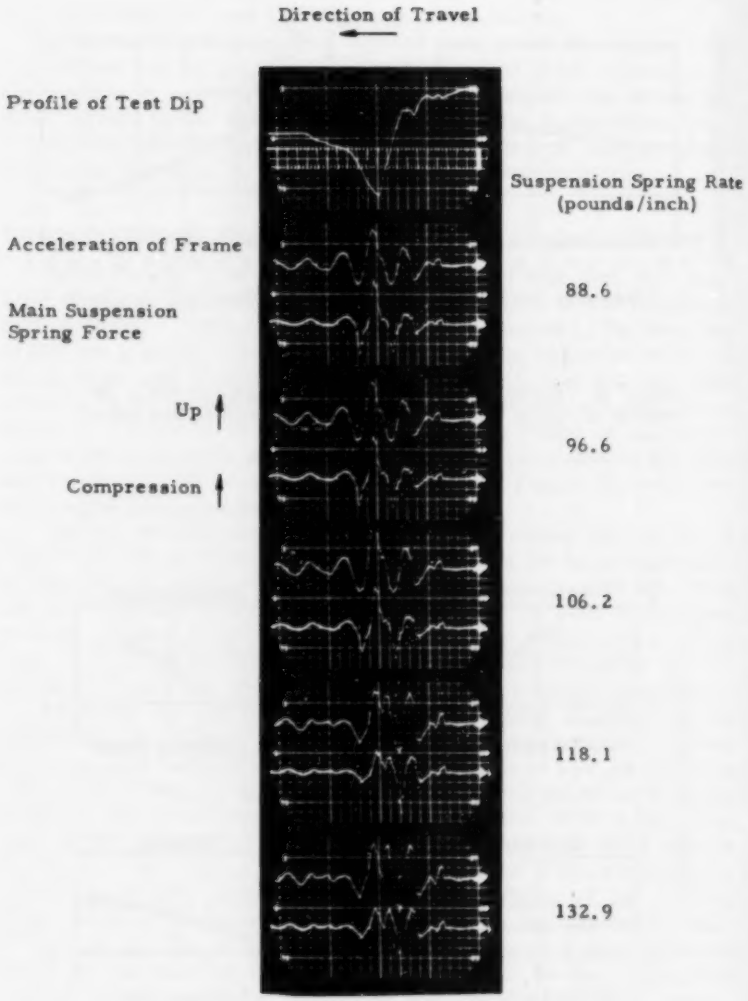


Figure 25

Effect of Varying the Main Suspension Spring Rate on the Motion at Front Wheels in Going Over Test Dip at 25 Miles Per Hour.

Wheel Bounce

Records obtained for motions over the half-sine bumps are reproduced in Figures 27 to 31. Measurements chosen for comparison are the wheel accelerations and the displacements of the wheels relative to the frame. In Figures 27 to 30 for symmetric motion, computed results are shown for both linear and non-linear systems. The values of viscous damping used in the linear cases were 44.0 slug/sec/wheel and 90.5 slug/sec/wheel at the front and rear, respectively. Non-linearities included in the computed results were the effect of tires leaving the ground and suspension damping (Figures 16 and 17). The actual vehicle measurements plotted for symmetric motion are the averages of those obtained on left and right hand sides of the vehicle. While the agreement is not as good as was obtained for the dip, it is still quite good considering the possible errors involved. In addition to the sources of error mentioned previously, here the tire was assumed to contact the ground at a point and the suspension bump stops were neglected in obtaining the analytical results.

A typical computer record of motion of the vehicle is shown in Figure 32. It can be clearly seen that the coulomb friction in the suspension gives rise to high rates of change of acceleration of the frame. A passenger in the vehicle commonly refers to the quality of the resultant ride as "jerky." Quantitative measurements of the magnitude of this effect could be made from computer records such as this. The damping force curve which has greater peaks in rebound reflects the unsymmetrical shock absorber characteristic.

No significant change in the primary motion over the test dip was found to take place with variation in the tire spring rate (tire pressure). However, varying the tire spring rate did have a strong effect on the peak accelerations experienced by the wheels as shown by the computed results in Figure 33. Higher frequencies of secondary motion were also observed with increasing tire spring rate, as would be expected.

Figure 31 shows results obtained for motion of the rear axle going over the wooden bump with the left wheels only. Analytical results are presented for a linear system only. It may be concluded from the results for symmetric motion that the inclusion of the non-linear damping and the effect of the tire leaving the ground in the system used for studying the rear axle motion would yield results showing agreement similar to that of Figure 30.

Tire Damping

The effect of the hysteresis of the tire, exhibited by the static load-deflection curves in Figure 18, was approximated in the

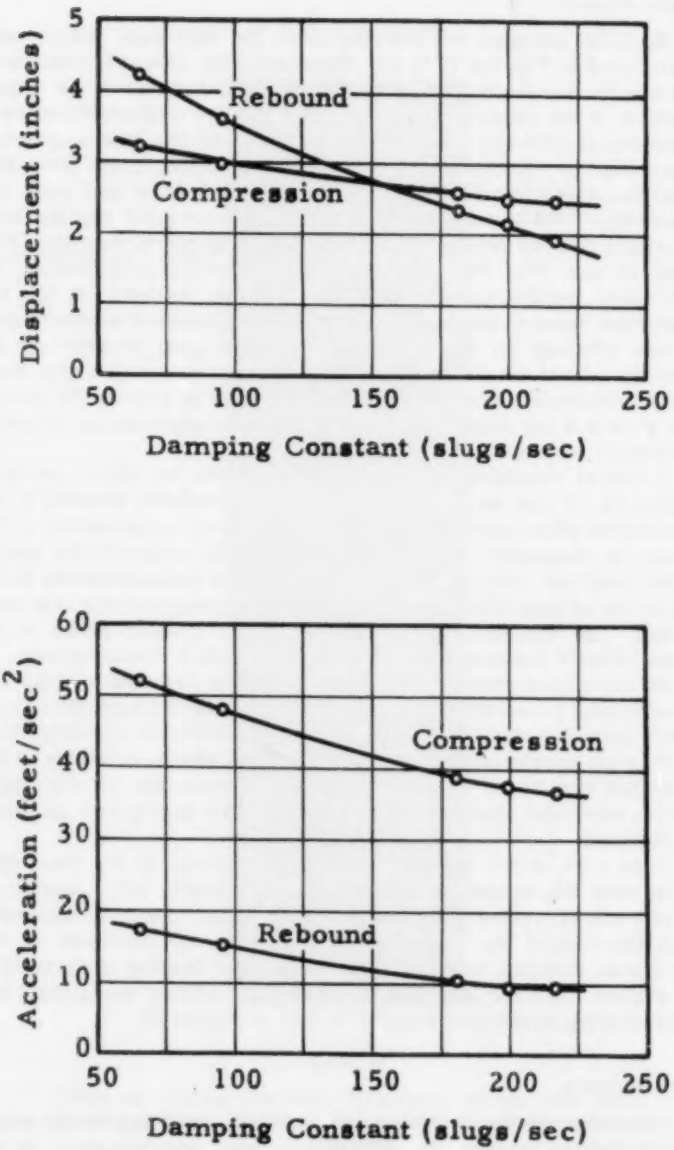


Figure 26

Effect of Varying Viscous Suspension Damping at the Rear Wheels in Going Over the Test Dip at 25 Miles Per Hour.

analysis by adding viscous damping, G_2, to the wheel. Values used for this damping were determined by trial and error. Computer results using different values were compared with vehicle test records until the best agreement was obtained. It was found that the values thus determined varied considerably with speed of travel and, consequently, with the amplitudes of the motion.

If one looks at the tire load-deflection curve for a tire pressure of 24 psi gage and assumes that the hysteresis loop is due to coulomb type friction (i.e., the width of the hysteresis loop remains constant regardless of amplitude), the magnitude of this friction force may be estimated as one half of the difference in load across the loop at a given deflection. The static deflection of the tire was approximately one inch; the coulomb friction force estimated from this portion of the curve is about $\pm$ 40 pounds. The equivalent viscous friction [18] g which will effect the same dissipation of energy as coulomb friction $\pm F$ for sinusoidal motion with an amplitude A and a frequency w is given by

$$g = 4F/\pi Aw.$$

This equation may be derived by calculating the amount of energy dissipated by each form of damping in one cycle and equating these energies. For motions of the unsprung mass, the natural frequency may be calculated from the vehicle test records. From the acceleration records in Figure 10, the period of this motion is about 0.11 second so that $w = 2\pi/0.11 = 57$ radians/second. The amplitude of the relative motion between the wheels and the frame for the first cycle of vibration after going over the half-sine bumps may be used for the amplitude A. This amplitude may be read from the graphs of the vehicle records. Using values of A obtained from Figures 27 to 29, and the values of F and w given above, the values of g given in Table 1 were calculated from the above equation. Also listed are the values of tire damping used on the computer.

TABLE 1

Comparison of Calculated and Experimental Values of Tire Equivalent Viscous Damping

Speed	Amplitude A	Equivalent Viscous Damping g	
(miles/hour)	(inches)	Calculated (slug/sec)	Computer (slug/sec)
10	0.2	54	57.5
15	0.75	14	20.4
20	1.4	7.7	2.6

146786

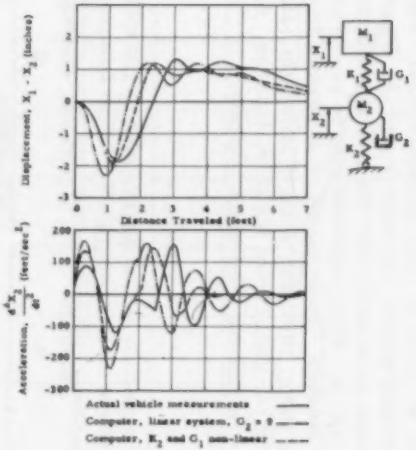


Figure 27

Motion at Front Wheels in Going Over Symmetric Half-Sine Bumps at 10 Miles Per Hour.

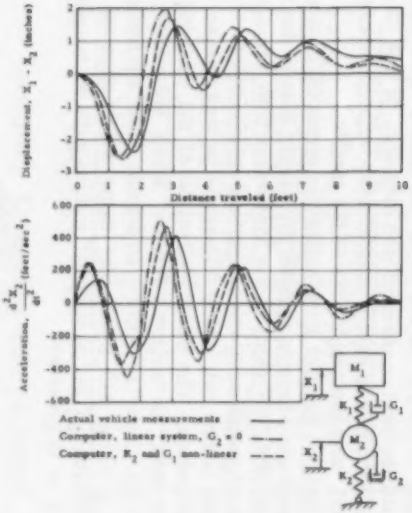


Figure 28

Motion at Front Wheels in Going Over Symmetric Half-Sine Bumps at 15 Miles Per Hour.

Half-Sine Bumps at 10 Miles Per Hour.

Half-Sine Bumps at 15 Miles Per Hour.

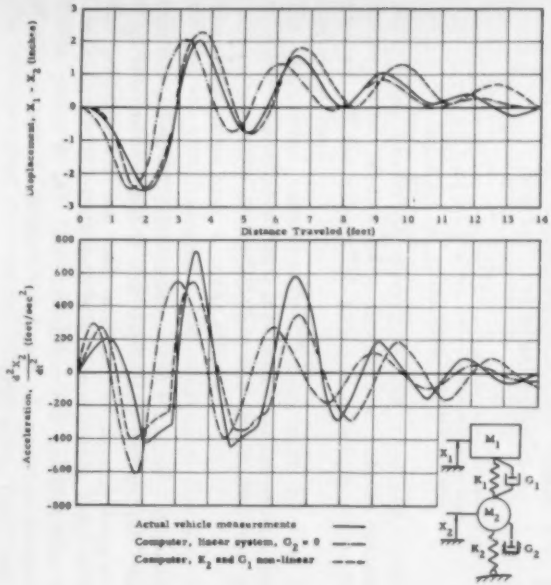


Figure 29

Motion at Front Wheels in Going Over Symmetric Half-Sine Bumps at 20 Miles Per Hour.

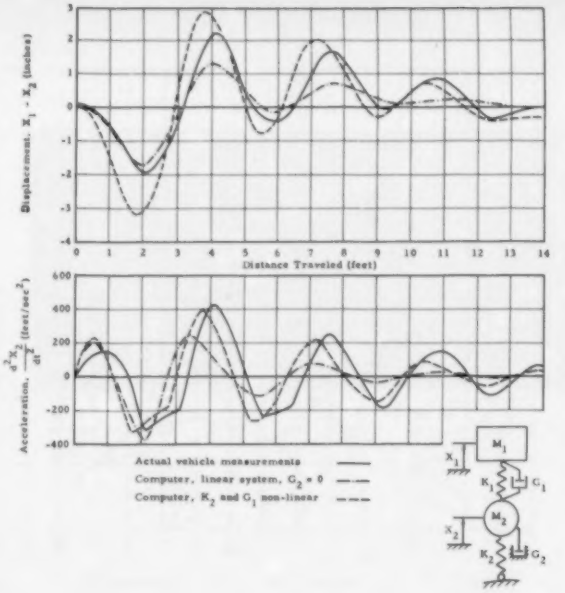


Figure 30

Motion at Rear Wheels in Going Over Symmetric Half-Sine Bumps at 20 Miles Per Hour.

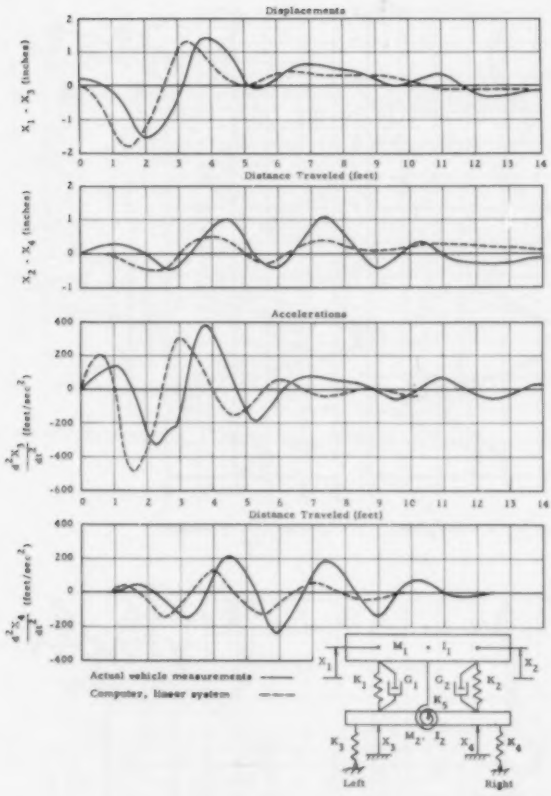


Figure 31

Motion of Rear Suspension in Going Over Half-Sine Bump With Left Wheels Only at 20 Miles Per Hour.

Considering the rather crude assumptions involved, the agreement between the values calculated and those determined by trial and error is good enough to conclude that the tire damping can be approximated from measurements of the static friction. Also, the agreement substantiates the statement that the damping added to obtain agreement between the computer results and the vehicle tests should be attributed to the tire.

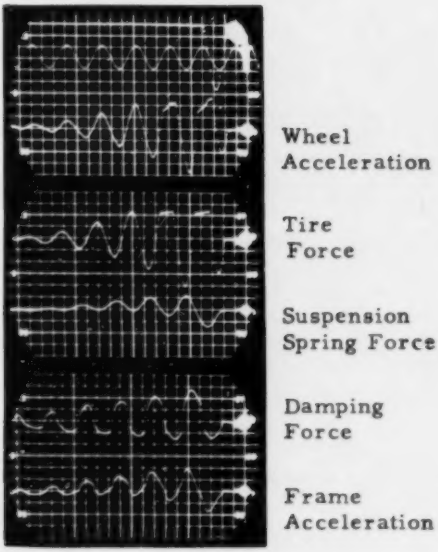


Figure 32

Computer Record Showing Motion at Front Wheels, Over Half-Sine Bumps at 20 Miles Per Hour.

CONCLUSIONS

During the course of conducting the research program from which the above results were drawn, the author reached certain conclusions on the effects of some automobile properties, which may be of interest.

1. The maximum accelerations experienced by the frame are strongly influenced by the presence of the suspension bump stops. Decreasing bump stop clearance and increasing the stiffness of these stops both have adverse effects.
2. Neither the non-linearity of the suspension damping characteristics nor the unsymmetrical action of the shock absorbers have significant effects on low frequency (about 1 cycle/sec) motion of the sprung mass.
3. Coulomb friction in the suspension system has a considerable effect on the accelerations of the frame caused by motion of the wheels at higher frequencies.
4. The incorporation of damping between the wheels and the frame

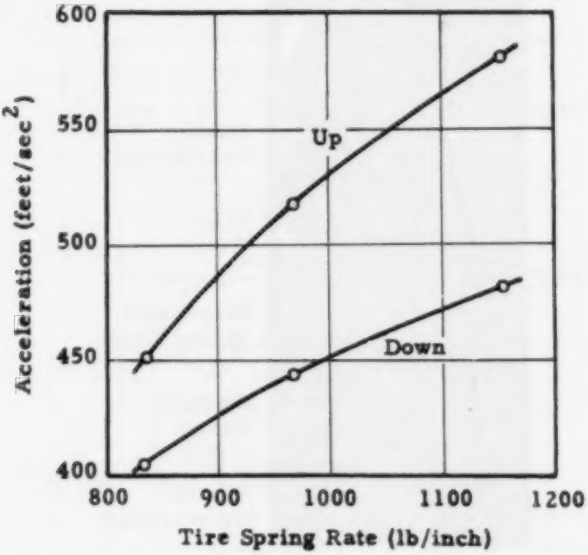


Figure 33

Effect of Varying the Tire Spring Rate at the Front Wheels, Half-Sine Bumps, 15 Miles Per Hour.

is wrong in principle. Ideally, the absolute motions of the sprung and unsprung masses should be damped independently. The practice of damping the relative motions between these masses is a compromise, at best. As a result, there is a wide variation of opinion about what the characteristics of shock absorbers should be.

5. Damping in the tire appears to have the characteristics of coulomb type friction. Therefore, the effect of this damping is greatest for small amplitudes of wheel motion, and the effect decreases as the amplitudes increase. Whether or not this damping may be neglected depends on the road excitation.

Of greatest importance is the following general conclusion:

Relatively simple systems can be used to represent the vehicle in motion over a variety of courses. These simple systems can be used to study the effects of varying vehicle properties. The results obtained are of sufficient accuracy to be employed for design purposes. It is possible to study the effects of different

properties (spring rates, damping characteristics, masses) separately. The practically unlimited number of variations make studies possible which will yield a clearer picture of the phenomena taking place.

ACKNOWLEDGEMENT

The author is indebted to Dr. Peter Kyropoulos for his great help in organizing and coordinating the research program, to Dr. Donald E. Hudson for his many suggestions and constructive criticisms, and to Dr. G. D. McCann for his advice on computers, all of the California Institute of Technology. Grateful acknowledgment is also made to the Research Laboratories Division of the General Motors Corporation, which furnished the test car, to Mr. Maurice Olley, Chevrolet Division of the General Motors Corporation, for furnishing data on the car and to the Caltech Mechanical Engineering Shop personnel for their work in outfitting the test car. Thanks also are due to Mr. Earl Phillips, of the Armour Research Foundation of the Illinois Institute of Technology, who assisted in editing this paper.

References

1. James J. Guest, "The Free Vibrations of an Autocar," *Engineering*, 120 (1925), 367.
2. Rudolf Slaby, "Einfluss der Achsmasse auf die Federungseigenschaften eines Fahrzeuges," *Automobiltechnische Zeitschrift*, 39 (1936), 593-598.
3. R. Schilling, and H. O. Fuchs, "Modern Passenger-Car Ride Characteristics," *Journal of Applied Mechanics*, 8 (1941), A59-A66.
4. A. Lange, and J. Dute, "Electronic Simulation as an Aid to the Design of Vehicle Suspension Systems," *SAE Preprint* No. 753 (1952).
5. Horace Grover, "Analog Helps Solve Suspension Problems," *SAE Journal*, 63 (April, 1955), 44-46.
6. A. W. Colwell, "Mobile Instrumentation for Automobile Equipment," *Proceedings of the Society for Experimental Stress Analysis*, 10 (1952), 65-76.
7. A. S. Bassette, "Instrumentation of Vibration in Automobile Body," *SAE Journal*, 59 (November, 1951), 37-39.

8. Robert J. Saxon, "Measuring for Car Shake," *SAE Preprint* No. 274 (1954).

9. M. Ruegg, "Laboratory Simulation of Car Shake," *SAE Preprint* No. 275 (1954).

10. P. E. Mercier, "Theory of Vehicle Suspension," *Journal of Applied Physics*, 13 (1942), 484-495.

11. R. D. Jeska, "A Comparison of Real and Simulated Automobile Suspension Systems," University of Michigan Report UMM-117 (February, 1953).

12. H. M. Jacklin and G. J. Liddell, "Riding Comfort Analysis," *Engineering Bulletin, Purdue University*, 17 (1933), 51.

13. G. D. McCann and H. E. Criner, "Mechanical Problems Solved Electrically," *Westinghouse Engineer*, 6 (1945), 49-56.

14. G. D. McCann, C. H. Wilts, and B. N. Locanthi, "Application of the California Institute of Technology Electric Analog Computer to Non-Linear Mechanics and Servo-mechanisms," *Transactions of the American Institute of Electrical Engineers*, 68 (1949), 652-660.

15. G. Korn and T. Korn, *Electronic Analog Computers*. New York: McGraw-Hill Book Co., Inc., 1952.

16. V. Bush and S. H. Caldwell, "A New Type of Differential Analyzer," *Journal of the Franklin Institute*, 240 (1945), 255-256.

17. Fred Gruenberger, *Computing Manual*. Madison: University of Wisconsin Press, 1953.

18. G. W. Housner and D. E. Hudson, *Applied Mechanics-Dynamics*. New York: D. Van Nostrand, 1950, p. 96.

7

Improving the Validity of Transducer Measurements with Computing Circuits

By S. A. Lippman
Tire Division U. S. Rubber Company

The meaning of the word "transducer" has undergone a change in recent years. Formerly transducers were devices for converting energy from one form to another. Now they are also instruments for creating variations in one physical quantity equivalent to variations in another. The secondary variation is usually generated for reasons of convenience. For instance, in the study of sounds, a transducer called a microphone produces electrical signals which correspond to accoustic variations in air pressure. If we wish to record the character of the sounds or to analyze frequency spectra, we will find that the operations are easier with the electrical signals than with the pressure variations. In contrast to the earlier definition, some transducers utilize sources of energy other than the primary variable. The goal is often to attain changes in some physical quantity at a large power level, while drawing very little power from the system under observation.

In this paper we will be concerned with the use of transducers for translating mechanical variables into electrical signals. Velocity, acceleration, pressure and displacement are examples of some of the primary variables. In all cases the conversion to electrical signals requires the displacement of some sensing element. In piezo-electric detectors there is the distortion of a crystal; in inductive pickups there is a moving armature or coil; and in strain gages there is the extension of the wires and motion of the supporting structure. Similar motions occur in differential transformers, potentiometers, variable capacitors, magnetostrictive heads and others. The requirement of a moving mass in these transducers results in limitations in their use and distortions in the process of conversion. A somewhat general method for coping with these limitations and distortions will be the theme of this paper. However, before undertaking this phase of the presentation, let us consider an elaboration of the meaning of "transducer."

Transducers are usually employed to study the behavior of a system when it is reacting to some kind of excitation and are also used to study the excitation that produces a known response. An accelerometer, for example, may be used to evaluate the forces acting on a vehicle. In this instance the output of the transducer is

not only determined by the variable being studied (the forces which act on the vehicle) and the properties of the accelerometer, but also by the system to which the device is attached (the vehicle). In a sense the entire system including the transducer itself may be treated as a more complex single transducer. Regardless of the terms we use, however, we must account for the influences of the transducer and the system being studied in order to have valid measurements.

The basic idea of this technique requires knowing the differential equation of detection by the overall transducer. Knowledge of the coefficients, however, is not required. The differential equation is expressed in terms of the variable being investigated and the output signal. A circuit is then constructed that simulates this differential equation. Next, the circuit and the transducer are matched to each other in a test procedure. We now apply the output of the transducer to this circuit and produce a new signal which is equivalent to the variable being studied. Let us illustrate the process by means of a concrete example.

The diagram of Figure 1 indicates the basic mechanism of a typical seismic transducer. This instrument may be used for the detection of displacements, velocities, or accelerations. The sensing unit shown in the diagram may be a strain gage, a slide wire, a reluctance bridge, a differential transformer, etc. The entire instrument moves in the process of its operation. An internal mass, **M**, is suspended from the case of the instrument by means of some flexible member. As the case moves, inertia tends to keep the mass at a constant position or at a constant velocity. This tendency, however, is compromised by the flexible member. In addition, there may also be a viscous drag which acts between the moving mass and the case. This drag is often purposefully included to minimize oscillations but may be due to mechanical hysteresis of the flexible member or to the "radiation" of vibratory energy out of the transducer to the system being measured. Finally, the operation of the transducer also depends on the sensing unit. This device detects the displacement of the mass from its position of rest in the case or in some instances it detects a property of the relative motion.

When we consider the quantitative relationships in the transducer, we obtain the following differential equation:

$$M(\ddot{X} + \ddot{Y}) + KY + \mu \dot{Y} = 0, \tag{1}$$

where **X** is the displacement of the case, **Y** is the displacement of the sprung mass (**M**) relative to the case, **K** is the spring constant of the flexible member, and μ is the coefficient of the viscous drag.

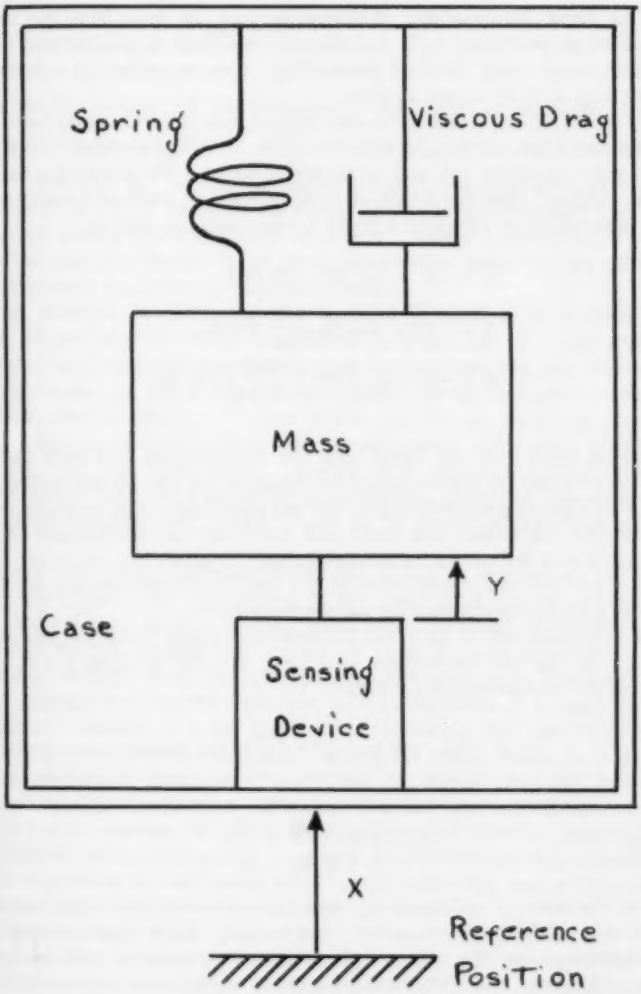


Figure 1

Schematic Diagram of Seismic Transducer

The transducer moves an amount X from the reference position. The corresponding movement of the sprung mass from its internal reference is Y. The sensing device detects some property of the variation Y.

Let us first assume that the sensing unit is such that the output signal is proportional to Y and that the constant of proportionality is 1. With very little loss in generality, then equation (1) relates the output signal to the input motion.

In order to see how this one device can act variously as a displacement gage, velocity detector and an accelerometer, we will rearrange equation (1) and integrate it twice. We obtain the following equations. The constants of integration are omitted since we are only interested in variations from a reference condition.

$$\ddot{X} = -(\ddot{Y} + \dot{Y}\mu/M + YK/M) \tag{2}$$

$$\dot{X} = -(\dot{Y} + Y\mu/M + \frac{K}{M}\int Y dt) \tag{3}$$

$$X = -(Y + \frac{\mu}{M}\int Y dt + \frac{K}{M}\iint Y dt\, dt) \tag{4}$$

When the nature of the input and the coefficients are such that the first two terms in the brackets of equation (2) can be neglected, the output is almost proportional to the acceleration. For another set of values and functions, the first and third terms in equation (3) are negligible and the output is almost proportional to the velocity. Similarly, we can see from equation (4) that there are realms in which the transducer approximates the displacement.

It may also be of interest to note the exact form of the output signal. If the accelerometer has been at rest at time t = 0, then integration of equation (1) yields:

$$Y = -\frac{1}{C}\int_0^t [\ddot{X}]\,(t-\tau)e^{-\mu\tau/2M}\,\mathrm{Sin}\,C\tau\,d\tau, \tag{5}$$

where

$$C = 1/2\sqrt{\frac{4K}{M} - \frac{\mu^2}{M^2}}$$

τ is a variable of integration, and the brackets indicate that X is taken as a function of $t-\tau$ during integration. This relation may also be expressed as the sum of the desired derivative and an error term. However, for reasons of brevity I have not included these forms. The relations of equation (5) are, in general, too complex for easy discernment of the details of X from the output signal.

Now let us look briefly at the frequency response of the transducer. This is a common way to evaluate the performance.

For a sinusoidal input at an angular frequency ω, the ratio of input to output is termed the transfer function $T(i\omega)$ and is given by

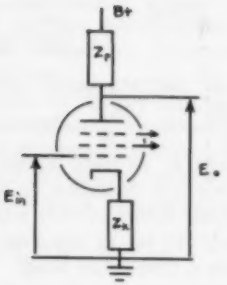
(6) $$T(i\omega) = \frac{1}{1 + \frac{\mu i}{\omega M} + \frac{K}{M\omega^2}}$$

In different ranges of performance, the transfer function can be almost equal to 1, $\frac{M\omega}{\mu}$, or $-\frac{M\omega^2}{K}$, depending on the relative values of ω, , M and K. Again, in each range the device acts as a displacement gage, a velocity detector or an accelerometer. In any frequency range the transducer is only somewhat correct in its conversions and is likely to be in considerable error if the subject matter covers too wide a frequency band.

The solution of restoring the desired derivative of **X** from the output of the transducer is contained in equations (2), (3) and (4). We have only to form the appropriate derivatives and integrals of Y and add these in the correct proportions. Graphical and numerical methods are possible, but are cumbersome as well as inaccurate. The use of an electrical analog computer at the time of measurement, on the other hand, results in a continuous flow of corrected data. The next point to be covered, then, is the nature of electrical analogs that might prove to be satisfactory for our purposes.

There are many approaches to the design of the electrical analogs. We might avail ourselves of the transconductance properties of vacuum tubes. In this way we can cause currents which are proportional to the transducer output to flow through networks. The resulting voltage drop is equivalent to the computed function. We might employ the current transfer characteristics of transistors in a similar manner, and we might even resort to the use of passive circuits for restricted frequency ranges. Finally, there is also the use of feed-back circuits—the method currently popular in commercial analog computers. Figures 2 and 3 contain schematic diagrams for a number of these approaches and indicate some of the quantitative relationships. The choice of a particular type of circuit is not important to our discussion. In all cases the use of the analog computer involves the same subsequent steps.

The constants of the transducer are presumably unknown, and in designing the circuits we do not usually match the values of the circuit elements to the exact requirements. Instead, adjustments are provided so that a sufficient number of parameters can be varied at will. The scheme of matching the computer to the transducer is based on a test procedure which taxes the accuracy of the computer. The input must be a known function of time, and incorrectly computed outputs should be readily recognized. The input might be step functions or sharp pulses (delta functions). The response of a poorly



Z_k	Z_p	$\Delta E_o \approx$
R_k	L C R CCD	$-\left[\frac{1}{C R_k}\int E_{in}dt + \frac{R}{R_k} E_{in} + \frac{L}{R_k}\dot{E}_{in}\right]$
L_k	L C R CCD	$-\left[\frac{1}{C L_k}\iint E_{in}dtdt + \frac{R}{L_k}\int E_{in}dt + \frac{L}{L_k} E_{in}\right]$
C_k CCD	L C R CCD	$-\left[\frac{C_k}{C} E_{in} + RC_k \dot{E}_{in} + L\ddot{E}_{in}\right]$
C_k L_k R_k	R	$-\left[C_k R \dot{E}_{in} + \frac{R}{R_k} E_{in} + \frac{R}{L_k}\int E_{in}dt\right]$
C_k L_k R_k	L	$-\left[C_k L \ddot{E}_{in} + \frac{L}{R_k}\dot{E}_{in} + \frac{L}{L_k} E_{in}\right]$
C_k L_k R_k	C CCD	$-\left[\frac{C_k}{C} E_{in} + \frac{1}{C R_k}\int E_{in}dt + \frac{1}{C L_k}\iint E_{in}dtdt\right]$

Figure 2

Computing Circuits Utilizing the Property of a Vacuum Tube

The A.C. impedance in the cathode circuit is great enough to enable the tube to act as a cathode follower. When series capacitors are employed, we use constant current devices (CCD) to provide the operating potentials to the tube. In principle the CCD does not affect the voltage drops across R,C,C_k and L or the currents through them. Pentode circuits, vacuum photo cells, and transistor circuits can be used for this purpose.

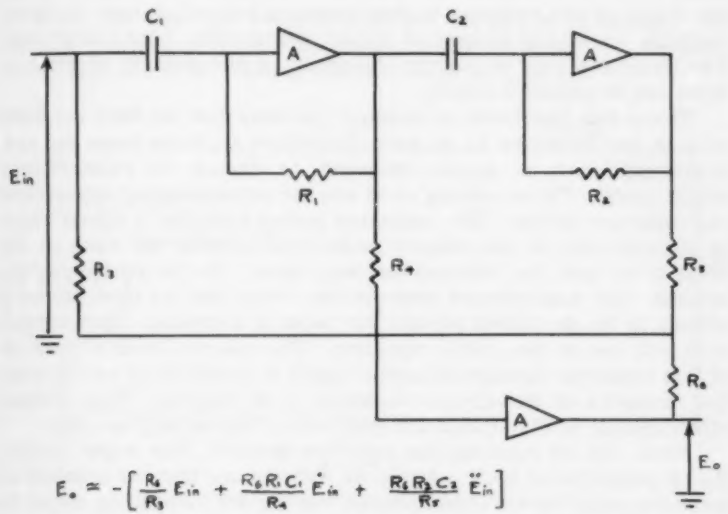


Figure 3

Typical R-C Feed-back Computer

The A's are high-gain inverting amplifiers. Similar circuits are possible for the other integro-differential equations. Successive differentiation by this method is to be avoided if possible. High-frequency noise becomes emphasized and very low noise levels are necessary in the power supplies and the transducer.

damped accelerometer to a step function in force, for example, would be a jump superimposed on a decaying sine wave. We can excite this response experimentally by applying a constant force through a string and cutting the string. In a series of repeated tests, the computer is adjusted until it converts this oscillatory signal into a step function with a very sharp rise time.

More elaborate adjustment procedures are also possible and may be preferred when the appropriate input functions are not practical or when the required settings are numerous. We can then establish the relative values of the parameters of the transducer by means of the response to sine waves of various frequencies. This procedure has been developed for the analysis of servomechanisms

and other linear systems. There is also an easier approach for some kinds of transducers. The required data are available from the detail of the response to step functions and unit pulses. In these methods we derive numerical values descriptive of the transducer. The computer can, therefore, be designed accordingly or the circuits can be adjusted exactly.

There are two kinds of sensing elements that we have not dealt with in our examples of seismic transducers. Since these two are commonly used, it appears desirable to discuss the relationships which apply. These sensing units are the piezo-electric crystal and the inductive pickup. The inductive pickup produces a signal which is proportional to the velocity difference between the case of the transducer and its internal seismic mass. In the piezo-electric method, the displacement between the mass and the case causes a charge to be developed across the faces of a crystal. This crystal also acts as an electrical capacitor. The charge tends to leak off of this capacitor through accidental paths of conductivity and through the resistors of the circuits to which it is coupled. This leakage effect results in distortions for motions that are slowly varying.

First, let us consider the inductive devices. The output voltage E_O is proportional to Y. Again we will assume that the constant of proportionality is 1. Consequently, the equations relating output to input are

$$\dddot{X} = -(\ddot{E} + \frac{\mu}{M}\dot{E} + \frac{K}{M}E), \tag{7}$$

$$\ddot{X} = -(\dot{E} + \frac{\mu}{M}E + \frac{K}{M}\int E dt), \text{ and} \tag{8}$$

$$\dot{X} = -(E + \frac{\mu}{M}\int E dt + \frac{K}{M}\iint E dt\, dt. \tag{9}$$

The third derivative of the displacement (jerk) is included because it can be approximated directly by the inductive pickup. On the other hand, the displacement itself cannot be approximated directly and is not listed. The conditions of the approximations are similar to that which holds for equations (2), (3) and (4). Equations (7), (8) and (9) are the relationships which are to be simulated in the analog computers when we attempt to reconstruct the original input to the inductive transducer.

The pertinent property of the piezo-electric crystal is that the rate of formation of charge is proportional to the rate of distortion. If q represents the charge, and as before Y represents the displacement, then:

(10) $$\dot{q} = \alpha \dot{Y}.$$

When we take into consideration the formation of change, the leakage resistance (r) and the capacity (C) of the crystal, we have the system which is shown in abstract form in Figure 4.

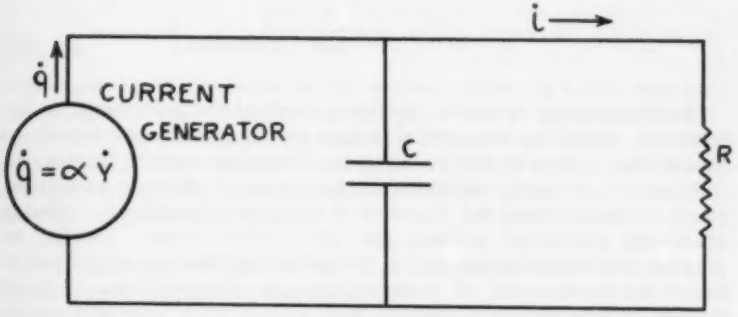


Figure 4

Equivalent Circuit Piezo-electric Transducer

The generator produces a current which is proportional to the rate of change of deformation. The charge tends to collect on the capacitor, but is partially bled off through the leakage paths, R.

According to Kirchoff's laws

(11) $$\dot{q} = -i + \alpha \dot{Y}_1, \text{ and}$$

(12) $$C\dot{E}_0 + \frac{E_0}{R} = \alpha \dot{Y}.$$

Once again we can let the coefficient of proportionality (α) be 1. We combine equation (12) with (2), (3) and (4) to determine the correcting equations for the piezo-electric transducer. The equation for the third derivative will also be included since it too may be approximated directly by the transducer for the proper combination of operating conditions and parameters.

(13) $$\dddot{X} = -\left[\dddot{CE} + \left(\frac{\mu}{M} + \frac{1}{R}\right)\ddot{E} + \left(\frac{KC}{M} + \frac{\mu}{MR}\right)\dot{E} + \frac{K}{MR}E\right]$$

(14) $$\ddot{X} = -\left[\ddot{CE} + \left(\frac{\mu}{M} + \frac{1}{R}\right)\dot{E} + \left(\frac{KC}{M} + \frac{\mu}{MR}\right)E + \frac{K}{MR}\int E dt\right]$$

$$(15)\quad \dot{X} = -\left[C\dot{E} + \left(\frac{\mu}{M} + \frac{1}{R}\right) E + \left(\frac{KC}{M} + \frac{\mu}{MR}\right) \int E\,dt + \frac{K}{MR} \iint E\,dt\,dt \right]$$

$$(16)\quad \dot{X} = -\left[CE + \left(\frac{\mu}{M} + \frac{1}{R}\right) \int E\,dt + \left(\frac{KC}{M} + \frac{\mu}{MR}\right) \iint E\,dt\,dt + \frac{K}{MR} \iiint E\,dt\,dt\,dt \right]$$

The computing circuits provide a method which in theory is capable of extending the useful range of transducers as faraas we would like. In practice we have to recognize certain limitations. First of all we rarely can determine an equation which is completely exact in representing the behavior of an actual transducer. What is more, the electrical analogs are not perfect either. Finally, we may have to compromise the entire effort by the use of a filter to lower the background of unwanted signals. Nevertheless, through the use of this technique, we have been able to extend our application of transducers to an appreciable extent.

8

Simplified Numerical Solution of a Catenary

By Elihu Geer
Professor of Civil Engineering
University of Detroit, Detroit

SYNOPSIS

Methods for the solution of the several cases of a catenary are presented, with examples, the only tables needed being those of common logarithms, and excellent accuracy being obtained with only one or two trials. The methods should be useful in solving problems arising in connection with tape measurements, cable conveyors, and electrical transmission lines.

The following difficulties are met in using the usual formulas for numerical solution of a catenary:

1. The solution involves cut and try, with the unknown both inside and outside a hyperbolic function or exponential.
2. The tables of hyperbolic functions which are commonly available are too coarse to yield accurate results.
3. Initial trial values are not apparent, and convergence may be slow at first.

There is offered below a system for the solution which has the following advantages:

1. In certain cases the initial trial value may be of satisfactory accuracy.
2. In many cases interpolation or extrapolation from a pair of trial values may yield a satisfactory answer.
3. In very unfavorable cases only a second interpolation may be needed.
4. No other tables than common or natural logarithms are used.

The notation for the simple case is shown in Figure 1.

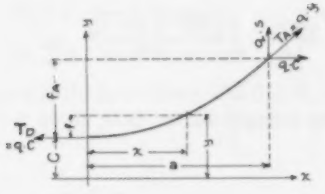


Figure 1.

C is defined as $\frac{T_D}{q}$, or $T_D = C.q$

q = load per unit length of cable

s = length of cable from low point to support

The derivation of formulas (some of which are well known) follows. From the condition that the cable is flexible and that consequently the slope of the cable equals the ratio of vertical to horizontal components of stress,

(1) $\frac{dy}{dx} = \frac{q.s}{q.C} = \frac{s}{C}$, $ds = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \cdot dx = \frac{1}{C}\sqrt{s^2 + C^2}\, dx$

(2) whence $s = \frac{C}{2}(e^{x/C} - e^{-x/C}) = C.\sinh \frac{x}{C}$

Since $\frac{dy}{dx} = \frac{s}{C} = \sinh \frac{x}{C}$, by integration:

(3) $y = C.\cosh \frac{x}{C}$

(4) Squaring and subtracting, $y^2 = s^2 + C^2$, or $y = \sqrt{s^2 + C^2}$.

But $y = f + C$ whence

(5) $y^2 = (f+C)^2 = s^2 + C^2$, or $C = \frac{1}{2f} \cdot (s^2 - f^2)$

Inverting $s + y = C\left(\sinh \frac{x}{C} + \cosh \frac{x}{C}\right) = Ce^{x/C}$,

(6) $\ln \frac{s+y}{C} = \frac{x}{C}$, or $x = C.\ln \frac{s+y}{C} = \frac{s^2 - f^2}{2f} \ln \frac{s+f}{s-f}$

or $x = \frac{(s+f)(s-f)}{2f} \cdot (2.302585) \cdot \left[\log (s+f) - \log (s-f)\right]$

The elastic elongation of the cable between low point and point of support =

$$\Delta s = \int \frac{T ds}{AE} = \int \frac{qy ds}{AE} = \int \frac{q\sqrt{s^2 + C^2}\, ds}{AE} \quad \text{or}$$

(7) $\Delta s = \frac{q}{2AE} \cdot \left[s.y + C.x\right]$

in which s, y, C are after elongation. A small amount of cut and try may be needed in this computation, but initial s should yield practically the final Δ s.

CASE 1. Given: s, f, q, A, E

Required: a, T_D, T_A, Δ s

Procedure: Solve (5) for C

Solve (6) for a

$T_D = q.C$

$T_A = q.y = q(f+C)$

Solve (7) for Δ s

Example 1. A tape which has a length of 100.000 feet under stress, is suspended with a sag of 1.000 foot below the ends.

$q = .017$ lb. ft.,

$A = .005$ sq. in.,

$E = 28 \times 10^6$ psi.

Required: Horizontal distance between end marks
Pull at ends
Elongation.

Solution: $C = \frac{(50+1)(50-1)}{2x1} = 1249.5$ ft.

$a = 1249.5 \ln \frac{50+1}{50-1} = 49.987$ ft.

Horizontal distance between ends = 99.974 ft.

$T_A = .017 (1 + 1249.5) = 21.26$ lb.

$$\text{Elongation} = \frac{.017}{2x.005x28x10^6} \cdot \left[(50x1250.5) + (1249.5x49.987)\right]$$

= .0076 ft. for half tape,

.0152 ft. for entire tape.

Evidently the tape should have been 100.011 ft. long in an unstressed condition to give a horizontal measurement of 100.000 feet with a sag of one foot.

CASE 2. Given: a, f, q

Required: s, T_D, T_A

Procedure: Solve for length of parabola of same span and sag, and consider approximately equal to that of catenary,

$$(8)\quad s \cong a + \frac{2}{3}a\left(\frac{f}{a}\right)^2 - \frac{2}{5}a\left(\frac{f}{a}\right)^4 + \frac{4}{7}a\left(\frac{f}{a}\right)^6 - \ldots$$

or

$$(9)\quad s \cong \frac{a}{2}\sqrt{1+\left(\frac{2f}{a}\right)^2} + \frac{a^2}{4f}\ln\left[\frac{2f}{a} + \sqrt{1+\left(\frac{2f}{a}\right)^2}\right]$$

Using this approximate s, solve (6) for x. In this computation, take two values of s to the nearest whole foot in the vicinity of the approximate value, and interpolate or extrapolate to determine the value of s which will yield x = a. In case extrapolation is needed, repeat the process with the improved value of s.

Then solve for C in (5)

$T_D = q.C$

$T_A = q.y = q(f+C)$

Example 2.

Given: $a = 90'$, $f = 60'$, $q = 0.9\#/\text{ft.}$, $A = .2647\square''$,

$E = 30\text{x}10^6$ psi

Required: s, T_D, T_A, elongation

Solution: $$s \cong \frac{90}{2}\sqrt{1+\left(\frac{2x60}{90}\right)^2} + \frac{(90)^2}{4x60}\ln\left[\frac{2x60}{90} + \sqrt{1+\left(\frac{2x60}{90}\right)^2}\right]$$

$$= 112.1 \text{ ft.}$$

for $s = 112$, $x = \frac{1}{2x60}\left(\overline{112}^2 - \overline{60}^2\right)\ln\frac{172}{52} = 89.1605$

for $s = 113$, $x = \frac{1}{2x60}\left(\overline{113}^2 - \overline{60}^2\right)\ln\frac{173}{53} = 90.3910$

Interpolating, $s = 112 + \frac{.8395}{1.2305} = 112.68$

$$C = \frac{1}{2x60}\left(\overline{112.68}^2 - \overline{60}^2\right) = 75.81 \text{ ft.}$$

$T_D = 68.23\#$, $T_A = 122.23\#$,

$$\Delta s = \frac{0.9}{2x.2647x30x10^6} \cdot \left[(112.68x135.81) + (75.81x90)\right]$$

$$= .00125' \text{ for half span.}$$

Example 3. Very deep catenary:

Given: 200′ span, 200′ sag.

$$s \cong 50 \sqrt{1+\left(\frac{2x200}{100}\right)^2} + \frac{(100)^2}{4x200} \ln\left[\frac{2x200}{100} + \sqrt{1+\left(\frac{2x200}{100}\right)^2}\right]$$

$$= 232.34 \text{ ft.}$$

$$\text{for } s = 232', \; x = \frac{\overline{232}^2 - \overline{200}^2}{2x200} \ln \frac{432}{32} = 89.74$$

$$\text{for } s = 233, \; x = \frac{\overline{233}^2 - \overline{200}^2}{2x200} \ln \frac{433}{33} = 91.96$$

$$\text{Improved } s = 232 + \frac{100-89.74}{91.96-89.74} = 236.62$$

$$\left.\begin{array}{l}\text{for } s = 236, \; x = 97.87 \\ \text{for } s = 237, \; x = 99.80\end{array}\right\} \text{Improved } s = 236 + \frac{100-97.87}{99.80-97.87} = 237.10'$$

$$\text{for } s = 238, \; x = 101.72, \; \text{Improved } s = 237 + \frac{100-99.80}{101.72-99.80}$$

$$= 237.10'$$

Even for such an extreme case, the third trial was unnecessary.

CASE 3. Given: s, a, q

Required: f, T_D, T_A

Procedure: From the first three terms of the expression for half of a parabola,

$$(10) \quad f \cong \frac{a}{\sqrt{6}} \sqrt{5 - \sqrt{115 - \frac{90s}{a}}}$$

Use this f as initial trial value in (6)

Solve for C in (5)

$T_D = q.C$, $T_A = q.y = q(f+C)$

Example 4. A wire 250 feet long is suspended from two points at the same level and 220 feet apart. Compute the sag.

Solution: $f \cong \frac{125}{\sqrt{6}} \sqrt{5 - \sqrt{115 - 90\,\frac{125}{110}}} = 61$ ft.

Try $f = 61$, $x = \frac{186 \times 64}{2 \times 61} \ln \frac{186}{64} = 104.098$

Try $f = 51$, $x = \frac{176 \times 74}{2 \times 51} \ln \frac{176}{74} = 110.630$

Improved $f = 51 + \frac{10 \times .630}{6.532} = 51.96$

Try $f = 52$, $x = \frac{177 \times 73}{2 \times 52} \ln \frac{177}{73} = 110.039$

Use $f = 52 + \frac{.039}{.591} = 52.066$

Try $f = 52.066$; $x = \frac{177.066 \times 72.934}{2 \times 52.066} \cdot \ln \frac{177.066}{72.934} = 110.00$; OK .

CASE 4. Given: s, q, T_A

Required: a, f

Procedure: $y = \frac{T_A}{q}$

$C = \sqrt{y^2 - s^2}$

$f = y - C$

Solve (6) for $a = x$

Example 5. A 100′ steel tape weighing .02 lb./ft. is given a 15 lb. pull. Compute the sag and the distance measured.

Solution: $y = \frac{15}{.02} = 750$ ft.

$C = \sqrt{(750)^2 - (50)^2} = 748.331$ ft.

$f = 750 - 748.331 = 1.669$ ft.

$a = x = C \cdot \ln \frac{s+y}{C} = 748.331 \ln \frac{800}{748.331} = 49.963$ ft.

Horizontal distance = 99.926 ft., neglecting stretch.

$\Delta s = \frac{.02 \times 3.4}{2x.02 \times 28 \times 10^6} \left[(50 \times 750) + (748.331 \times 49.926) \right] = .005$ ft. for half tape

Corrected s = 50.005 ft.

Corrected $y = 750 \times \frac{50.005}{50} = 750.075$ ft.

Corrected $C = \sqrt{(750.075)^2 - (50.005)^2} = 748.406$

Corrected $a = x = 748.406 \ln \frac{800.080}{748.406} = 49.968$ ft.

Corrected Horizontal distance = 99.936 ft.

CASE 5. The points of support are at different levels, as shown in Figure 2.

Figure 2.

Given: (a+b), f_A, f_B, q

Required: a, b, s_A, s_B, T_D, T_A, T_B

Procedure: From the properties of the parabola, $a \cong \frac{a+b}{1+\sqrt{\frac{f_B}{f_A}}}$

Compute s_A from (8) or (9).

Choose a whole number of feet as a trial value for s_A, and compute the corresponding value of s_B in $\frac{s_A^2 - f_A^2}{2f_A} = \frac{s_B^2 - f_B^2}{2f_B} = C$

Solve for span in (a+b) = C. $\ln \frac{(s_A + f_A)}{(s_A - f_A)} \cdot \frac{(s_B + f_B)}{(s_B - f_B)}$

Take another value of s_A and repeat the process.

Interpolate or extrapolate to obtain an improved value for s_A and repeat until the derived value of a+b satisfies the given condition.

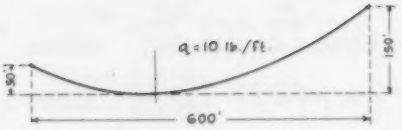


Figure 3.

Example 6. The problem is shown in Figure 3.

$$a \cong \frac{600}{1+\sqrt{\frac{50}{150}}} = 380.4 \text{ ft.}$$

$$s_A \cong 380.4 + \frac{2}{3}\,380.4\left(\frac{150}{380.4}\right)^2 - \frac{2}{5}\,380.4\left(\frac{150}{380.4}\right)^4 + \frac{4}{7}\,380.4\left(\frac{150}{380.4}\right)^6 - \ldots$$

$$= 416.7 \text{ ft.}$$

Try $s_A = 416$, $\dfrac{(416)^2-(150)^2}{2\text{x}150} = \dfrac{s_B^2-(50)^2}{2\text{x}50}$, $s_B \cong 229.5$, $C \cong 501.85$

$$a+b \stackrel{?}{=} 501.85 \cdot \ln\frac{566}{266} \cdot \frac{279.5}{179.5} = 601.18$$

Try $s_A = 415$, $\dfrac{(415)^2-(150)^2}{2\text{x}150} = \dfrac{s_B^2-(50)^2}{2\text{x}50}$, $s_B \cong 228.93$, $C \cong 499.08$

$$a+b \stackrel{?}{=} 499.08 \cdot \ln\frac{565}{265} \cdot \frac{278.93}{178.93} = 599.43$$

Improved value of $s_A = 415 + \dfrac{.57}{1.75} = 415.32$

Try $s_A = 415.32$ ft., $s_B = 229.12$, $C = 499.97$ ft.

$$a+b \stackrel{?}{=} 499.97 \cdot \ln\frac{565.32}{265.32} \cdot \frac{279.12}{179.12} = 599.984$$

$s_A = 415.33$, $s_B = 229.128$ $C = 499.997$ ft.

$$a+b \stackrel{?}{=} 499.997 \cdot \ln\frac{565.33}{265.33} \cdot \frac{279.128}{179.128} = 600 \text{ ft.; OK .}$$

CASE 6. Given: $(s_A+s_B) = M$, $(f_A-f_B) = K$, $(a+b) = L$

Required: a, b, f_A, s_A, s_B, f_B.

Procedure: Assume a value for s_A.

Compute the corresponding value of f_A and C in

$$C = \frac{s_A^2 - f_A^2}{2f_A} = \frac{(M-s_A)^2 - (f_A-K)^2}{2(f_A-K)}$$

Try these values of s_A, f_A, and C in

$$a+b = C.\ln \frac{(s_A+f_A)(M-s_A+f_A-K)}{(s_A-f_A)(M-s_A-f_A+K)}$$

Assume another value for s_A and repeat the above.

Interpolate or extrapolate to an improved value for s_A and repeat the process until a+b satisfies the given condition.

Compute remaining unknowns as before.

CASE 7. See Figures 4 and 5.

Given: q = weight per unit length of cable

W = concentrated load

K = difference in altitudes of supports

L = Total span

s'_A = length of cable from concentrated load to high support

s'_B = length of cable from concentrated load to low support

Required: Location of low point, tensions in cable.

Procedure: Due to equilibrium of horizontal forces,

(11) $$C = \frac{s_A^2 - f_A^2}{2f_A} = \frac{s_B^2 - f_B^2}{2f_B} = \frac{s_1^2 - f_1^2}{2f_1} = \frac{s_2^2 - f_2^2}{2f_2}$$

Since $(f_A - f_1) - (f_B - f_2) = K$,

(12) $$\sqrt{s_A^2 + C^2} - \sqrt{s_B^2 + C^2} - \sqrt{s_1^2 + C^2} + \sqrt{s_2^2 + C^2} = K$$

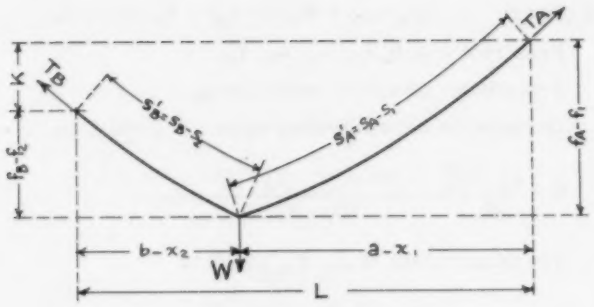


Figure 4.

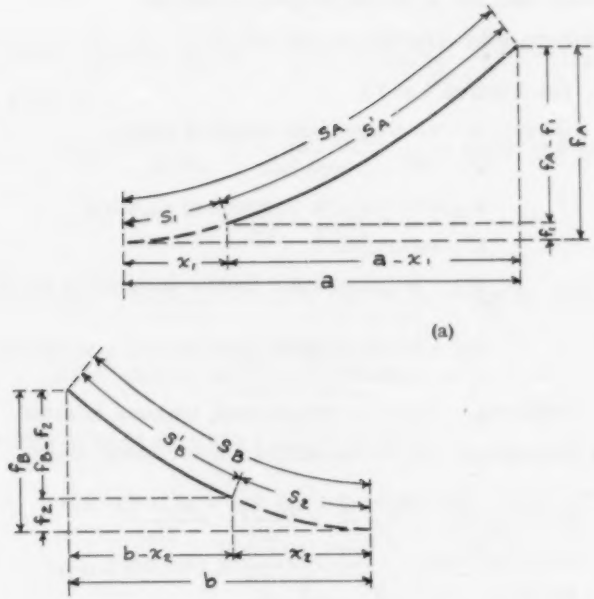


Figure 5.

Due to equilibrium of vertical forces,

$$(13)\quad s_1 + s_2 = \frac{W}{q}$$

$$(14)\quad s_A + s_B = (s'_A + s_1) + (s'_B + s_2) = s'_A + s'_B + \frac{W}{q}$$

Assume one of the four lengths and compute the others. A reasonable starting point is $s_1 = s_2 = \frac{W}{2q}$. In that case,

$$C = \sqrt{\left(\frac{s_A^2 + K^2 - s_B^2}{2K}\right)^2 - s_A^2}$$

Solve for C by cut and try in (12).

Then $f_A = \sqrt{s_A^2 + C^2} - C$, $\qquad f_B = \sqrt{s_B^2 + C^2} - C$

$f_1 = \sqrt{s_1^2 + C^2} - C$, $\qquad f_2 = \sqrt{s_2^2 + C^2} - C$

$$a = C \cdot \ln \frac{s_A + f_A}{s_A - f_A} \qquad b = C \cdot \ln \frac{s_B + f_B}{s_B - f_B}$$

$$x_1 = C \cdot \ln \frac{s_1 + f_1}{s_1 - f_1} \qquad x_2 = C \cdot \ln \frac{s_2 + f_2}{s_2 - f_2}$$

Then test to see if $a - x_1 + b - x_2 = L$.

If not, try another s, repeat, and interpolate.

The work in this test may be shortened somewhat by using:

$$(15)\qquad L = C \cdot \ln \frac{(s_A + f_A)(s_A + f_B)(s_1 - f_1)(s_2 - f_2)}{(s_A - f_A)(s_B - f_B)(s_1 + f_1)(s_2 + f_2)}$$

instead of computing a, b, x_1, x_2 individually.

Example 7. The problem is given in Figure 6.

Solution: $s_1 + s_2 = \frac{600}{10} = 60$. Assume $s_1 = s_2 = 30'$

Then $s_A = 209 + 30 = 239'$, $s_B = 105' + 30' = 135'$

$$C = \sqrt{\left(\frac{239^2 + 40^2 - 135^2}{2 \times 40}\right)^2 - 239^2} = 446.226$$

$f_1 = f_2 = 1.007'$ $\quad x_1 = x_2 = 29.968'$

$f_A = 59.974'$ $\quad a = 228.837'$

$f_B = 19.974'$ $\quad b = 133.123'$

$$a + b - x_1 - x_2 = 302.024'.$$

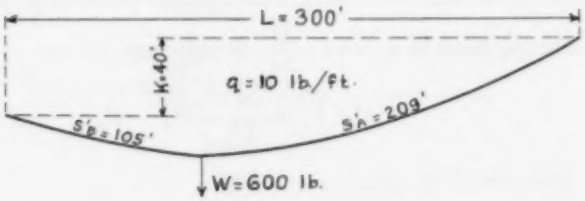


Figure 6.

A fair approximation for the location of the low point now can be found by assuming that the chord lengths from low point to points of support remain constant.

The horizontal location thus computed is only .04 foot in error, while the vertical location is approximately 0.9 ft. too low. There is no obvious choice for the next trial values for lengths. Trying $s_A = 235'$, $s_B = 139'$, $s_1 = 26'$, and $s_2 = 34'$, it is found that $C = 412.93'$, $f_A = 62.187'$, $f_B = 22.767'$, $f_1 = .817'$, $f_2 = 1.397'$, and that $a+b-x_1-x_2 = 300.457'$. Extrapolating, a value of slightly less than 234' for s_A is indicated. Trying $s_A = 234'$, $s_B = 140'$, $s_1 = 25'$, $s_2 = 35'$, it is found that $C = 404.560'$, $f_A = 62.799'$, $f_B = 23.539'$, $f_1 = 0.772'$, $f_2 = 1.511'$, $a = 222.597'$, $b = 137.346'$, $x_1 = 24.984'$, $x_2 = 34.957'$, which satisfy the conditions of the problem quite well. The 600 pound load is 102.39 feet from the left support measured horizontally, and 22.03 feet measured vertically. The tensions are:

$$T_A = 10(62.799+404.560) = 4673.59 \text{ lb.}$$

$$T_B = 10(23.539+404.560) = 4280.99 \text{ lb.}$$

9

The Mathematical Background of Spiroid Gears

By Oliver Saari, Development Engineer
Illinois Tool Works, Chicago

1. THE GENERAL FIRST-ORDER GEAR TOOTH ACTION LAW

Spiroid gears are more easily understood if some basic concepts regarding gear tooth action are first clarified. In fact, it was the formulation of these concepts which led to the development of the new gears.

Most gears rotate on fixed axes, at a constant angular velocity ratio. The axes can be parallel, intersecting, or skew, and many types of gears have been made for operation in each of these relative positions. Conjugate action of all of these gears can be shown to be in conformity with a perfectly general first-order gear tooth action, which will be here briefly outlined.

In general, active gear tooth faces are continuous surfaces having continuous first derivatives, i.e., they are smooth surfaces having no gaps or folds. They have instantaneous tangential contact at one or more points, which will be called "conjugate points." If the gear teeth have line contact, which is most often the case, there are an infinite number of these conjugate contact points for any instantaneous position of the gears, and the aggregate of these points makes up the contact line. The requirement that gears transmit motion between two fixed axes at a constant angular velocity ratio imposes certain restrictions on the tooth surfaces at every one of their conjugate points.

It is convenient to think of the contact at a conjugate point as being two infinitesimal particles, lying instantaneously in the common tangent plane. In general these particles have a relative motion, which results from the rotation of the gears on their axes. The fact that one gear tooth is imparting motion to the other implies that the direction of relative motion of the contacting particles must lie in the common tangent plane of the tooth surfaces, as illustrated in Figure 1.

The absolute velocities of the contacting particles are $\vec{V}_1$-$\vec{V}_2$. Their relative velocity is the difference, $\vec{V}_1$-$\vec{V}_2$. From the laws of vector subtraction, this is known to be a vector parallel to the line $\overline{ac}$ which joins the ends of $\vec{V}_1$ and $\vec{V}_2$. From study of Figure 1 can be inferred the fact that if the relative velocity lies in the tangent

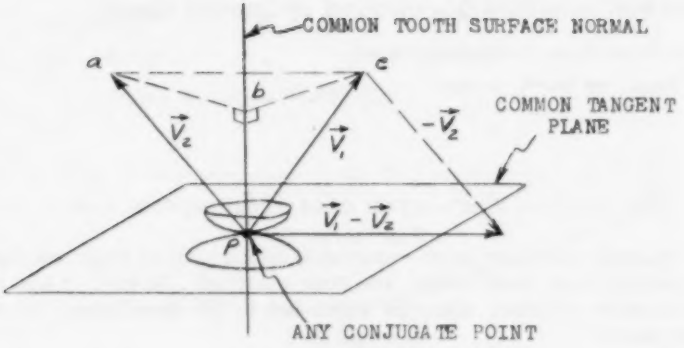


Figure 1

Basic Velocity Relationships

plane, then the normal components of the absolute velocities are equal, and conversely. This is true in any case where one surface drives another through a tangential contact point. This, and the fact that gears must operate on fixed axes at constant angular velocity ratio, are the conditions which determine the tooth-action law. This law is called "first-order" because it deals only with the tangency and common tangent plane of the tooth surfaces and says nothing about higher-order properties such as surface curvatures. A familiar first-order tooth action law is the two-dimensional law for transverse sections of parallel-axis gears: the common normal to mating tooth curves at every conjugate point passes through a fixed point, called the "pitch point." It is not surprising to find that the general first-order tooth-action law includes this as a special case.

The most general disposition of axes in a fixed frame of reference is illustrated in Figure 2.

For consistency in nomenclature, let us call the gear axes #1 and #2, with axis #1 always so placed that it coincides with the x-

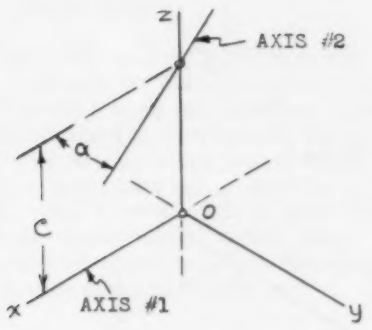


Figure 2

The Most General Disposition of Axes

axis of the Cartesian coordinate system. The angle between the axes is α and the shortest distance between them is C. The axes are parallel when $\alpha = 0$. They intersect, as in bevel gearing, when $C \neq 0$ but $\alpha \neq 0$. They are skew, as in worm, hypoid, and spiroid gears, when $C \neq 0$ and $\alpha \neq 0$.

This general case will be returned to later. For simplicity in demonstrating principles, let us first consider the right-angle skew-axis case, for which C is arbitrary and $\alpha = 90^\circ$, as shown in Figure 3. Moreover, let us restrict the analysis to a left-hand system, in which the angular velocities of the gears, $\vec{w}_1$ and $\vec{w}_2$, are directed along the axes as shown. Let $\underline{K}$ be the fixed angular speed ratio of the gears, and $\underline{p}$ (x,y,s) be any conjugate point. At every conjugate point $\underline{p}$, then, are two conjugate particles, each fixed to and rotating with its associated gear. The location vectors of these particles are $\vec{r}_1$ and $\vec{r}_2$ and their absolute velocities are $\vec{V}_1$ and $\vec{V}_2$.

Let

N_1 = number of teeth in gear #1

N_2 = number of teeth in gear #2

K = angular speed ratio = $N_2/N_1 = |\vec{w}_1|/|\vec{w}_2|$

Let $\vec{i}$, $\vec{j}$, and $\vec{k}$ be the usual unit vectors associated with the x, y, z axes. We may assume any convenient value for the angular speed of one gear, as long as the speed of the other is then made

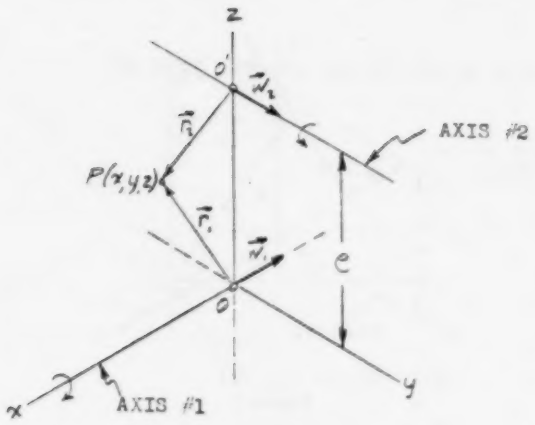


Figure 3

Right-Angle Skew-Axis Gear System

consistent with the speed ratio K. The following values are consistent:

$$\vec{w}_1 = -\vec{K}\vec{i}$$

$$\vec{w}_2 = \vec{j}$$

The location vectors of the conjugate particles can be seen from Figure 3 to be

$$\vec{r}_1 = x\vec{i} + y\vec{j} + z\vec{k}$$

$$\vec{r}_2 = \vec{r}_1 - C\vec{k}$$

Absolute velocities are given by cross-multiplication according to the rules of vector analysis,

$$\vec{V}_1 = \vec{w}_1 \times \vec{r}_1 = (Kz)\vec{j} - (Ky)\vec{k}$$

$$\vec{V}_2 = \vec{w}_2 \times \vec{r}_2 = -(C-z)\vec{i} - x\vec{k}$$

Let

$$\vec{V} = \text{relative velocity} = \vec{V}_1 - \vec{V}_2$$

Then

(1) $\vec{V} = (C-z)\vec{i} + (Kz)j + (x-Ky)\vec{k}$

The components of the relative velocity vector are functions of the coordinates. Obviously its direction at any point is not affected by speeding up or slowing down the rotations of the gears, as long as the angular speed ratio is maintained. Equation (1) defines a vector field of infinite extent, since obviously every point of space referred to the x, y, z coordinates has some value of relative velocity associated with it. Such a vector field is more easily visualized if we can discover its "field lines," or the family of space curves which are everywhere co-directional with the vectors of the field. The differential equation of the field lines is

$$dx : dy : dz = (C-z) : (Kz) : (x-Ky)$$

This can be more easily solved if coordinates are transformed as follows:

$$x = \frac{1}{\sqrt{1+K^2}}(Kx'-y')$$

$$y = \frac{1}{\sqrt{1+K^2}}(x'+Ky')$$

$$z = z' + \frac{1}{1+K^2}$$

This amounts to a rotation and translation of the axes to a new position shown in Figure 4.

The translation and rotation of the coordinate axes are determined from the speed ratio K, thus

$$\tan \beta = \frac{1}{K}$$

$$h = \frac{C}{1+K^2}$$

Let i', j', and k' be unit vectors corresponding to the new coordinate axes. The transformed equation of the relative velocity vector field becomes

(2) $$\vec{V} = \sqrt{1+K^2}\left[\left(\frac{KC}{1+K^2}\right)\vec{i'} + z'\vec{j'} - y'\vec{k'}\right]$$

and the differential equation of the field lines becomes

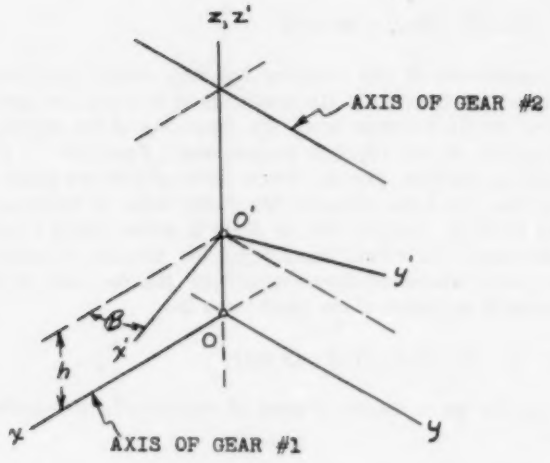


Figure 4

Transformation of Coordinates

$$dx' : dy' : dz' = \frac{KC}{1+K^2} : z' : -y'$$

This system of simultaneous differential equations can be easily solved, resulting in the equations of two surfaces whose intersecting curves are the field lines.

$$y'^2 + z'^2 = \text{an arbitrary constant} = \rho^2$$

$$x' - \left(\frac{KC}{1+K^2}\right) \sin^{-1}\left(\frac{y'}{\rho}\right) = \text{another arbitrary constant.}$$

The first of these equations represents a family of cylinders whose axes coincide with the x'-axis, and which have arbitrary radii ρ. The second equation represents a family of left-handed radial-helicoidal surfaces of equal lead, axes on the x' axis, and arbitrary phase relationship. The lead L_R is given by

$$L_R = \frac{2\pi KC}{1+K^2}$$

Simultaneously the two equations represent all left-handed helical space curves of the above lead, whose axes coincide with the x'-axis. These helices are the easily-visualized field lines of the relative velocity vector field. The x'-axis (Figure 4) will be called

the *relative axis* and the field lines the *relative helices*. For the right-angle gears, giving the center distance and speed ratio determines the position of the relative axis and the constant relative lead. For a given center distance and ratio, it is evident that through every point in the space passes one and only one relative helix. The relative axis is itself one of the relative helices, corresponding to $\rho = 0$.

These methods can be extended to the more complex case of general disposition of the axes, where α, C, and K are all arbitrary. In terms of these quantities the position of the relative axis and the value of the relative lead are

$$(3) \qquad \tan \beta = \frac{\sin a}{K+\cos \alpha}$$

$$(4) \qquad h = \frac{C(1+K \cos \alpha)}{1+K^2 + 2K\cos \alpha}$$

$$(5) \qquad L_R = \frac{2\pi KC \sin \alpha}{1+K^2 + 2K\cos \alpha}$$

These are illustrated in Figure 5.

The general first-order tooth-action law can now be stated as follows:

A necessary condition for conjugate action at any point is that the common tangent plane must also be tangent to the relative helix at the point.

That this is a perfectly general law, reducing in special cases to the well-known laws of conjugate action, can be readily demonstrated. For example, in parallel-axis gearing, since $\alpha = 0$,

$$\beta = 0$$

$$h = \frac{C}{1+K}$$

$$L_R = 0$$

Here the relative axis is a line parallel to the axes of the gears and co-planar with them. It is the tangent line of the pitch cylinders as conventionally defined. The relative helices have zero lead, being annular circles concentric with the relative axis; therefore in any transverse plane they appear as concentric circles with center at the pitch point. The general law is here equivalent to stating that the trace of the common tangent plane must be tangent to the relative circle passing through the conjugate point. But, as illustrated in

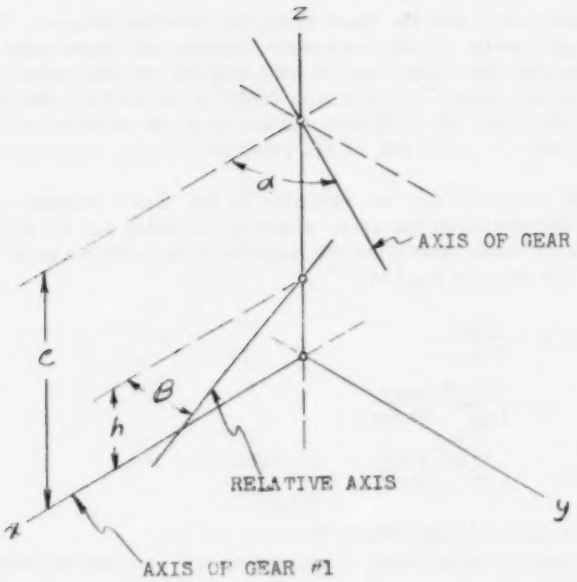


Figure 5

Relative Axis for General Case

Figure 6, this is exactly the same as saying that the common normal of the contacting tooth profiles must pass through the pitch point.

For intersecting-axis gears, $C = 0$ and $\alpha \neq 0$. Substituting these values in equations (3), (4), and (5) gives

$$\tan \beta = \frac{\sin \alpha}{K + \cos \alpha}$$

$$h = 0$$

$$L_R = 0$$

Here the relative axis is the contacting element of the well-known bevel gear pitch cones, and the relative helices are again annular circles about this line. Here the general law is seen to be equivalent to the statement that all common tooth normals at conjugate points must pass through the pitch line.

In these simpler cases, no apparent advantage results from restating the tooth-action law in terms of relative helices instead of having tooth normals intersect the pitch lines. It is in its applica-

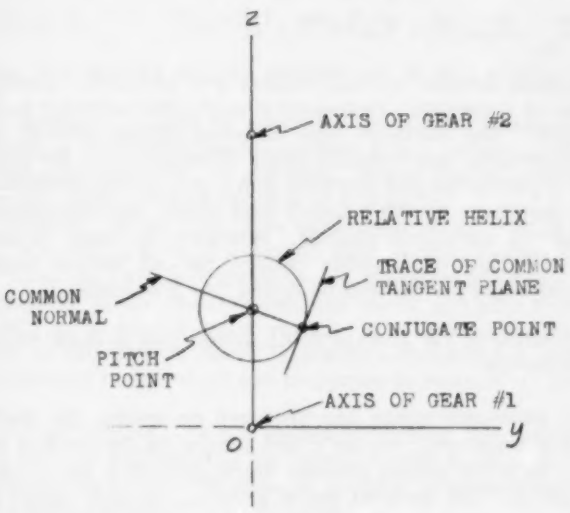


Figure 6

General Law as it Applies to Parallel Axes

tion to skew-axis gearing that the general law gives insight not anticipated by the conventional ways of stating the law. In skew-axis gearing the common tooth normals do not necessarily pass through the relative axis. At any conjugate point, the common tooth normal can lie anywhere in the plane normal to the relative helix at that point.

Certain interesting facts about the relative helices can be fairly easily proven. The amount of "sliding" action between gear teeth is constant alone any relative helix and is always directed along the helix. It is a minimum along the relative axis itself.

Moreover, any relative helix generates tangent surfaces of revolution about the axes of the gears. For example, if a given relative helix is used as a generatrix to describe a surface of revolution about axis #1, a surface normal vector is given by the formula

$$\vec{V}_{n_1} = \vec{V}_1 \times (\vec{V}_1 - \vec{V}_2) = - \vec{V}_1 \times \vec{V}_2$$

If the same relative helix is used as a generatrix to describe a surface of revolution about axis #2, the corresponding surface normal vector is given by

$$\vec{V}_{n_2} = \vec{V}_2 \times (\vec{V}_1 - \vec{V}_2) = -\vec{V}_1 \times \vec{V}_2$$

But these surface normal vectors are identical. Therefore the surfaces of revolution are tangent to each other at every point of the generatrix. The surfaces have constant sliding velocity at every point of contact, this velocity being directed along the generatrix. One such pair of tangent surfaces is the pair of hyperboloids of revolution generated by the relative axis itself, the well-known "pitch surfaces" of skew-axis gearing. However, surfaces of revolution generated by any of the other relative helices have the same properties and might just as well be called "pitch surfaces."

2. Application of the General Tooth Action Law to Skew-Axis Gear Design

Most skew-axis gears are produced by making the teeth of one member of some arbitrary form and generating the teeth of the other member by some cutting process which duplicates the rolling action of the gears. Conventional worm gearing, formate hypoid gearing, and spiroid gearing are of this type. In each case, the tooth form of the primary member is first chosen as one which can be made by some simple form-cutting process. A cutting tool is made whose form represents as nearly as possible the teeth of the primary gear member. This tool is then rotated with the work-blank of the conjugate gear member in a manner which duplicates the relative motion of the gears in action.

This method results in useful gears only if the tooth-form of the primary member is chosen with some care: it must be such that conjugate action is possible over a large portion of the available tooth surface. An analysis of possible contact lines in this type of gearing can be made by applying the general tooth action law as follows:

Let

$$F(x, y, z) = 0 \tag{6}$$

represent the pre-determined tooth form of gear #1 in some instantaneous angular position. Direction numbers of the tooth normal at any point are the partial derivatives F_x, F_y, and F_z. For conjugate action, the tooth normal must also be normal to the relative velocity vector. Therefore

$$(7) \qquad \phi(x, y, z) = 0 = V_x F_x + V_y F_y + V_z F_z$$

where V_x, V_y, and V_z are the components of the relative velocity vector V.

If the surfaces defined by equations (6) and (7) intersect, their intersection is the instantaneous contact line. If they do not intersect, conjugate action is not possible. Interesting general conclusions can be drawn from an analysis of this sort.

For example if the arbitrary tooth form of gear #1 is a plane surface, then $F(x, y, z)$ is of no higher than the first degree in any of the variables. Its partial derivatives are therefore constants. Since the components of the relative velocity vector are also of first degree, $\phi(x, y, z)$ is also a plane surface. The contact line is the intersection of two planes, which is always a straight line. This is a perfectly general result which can be proven in many more laborious ways.

The contact analysis is not always as easy as this, particularly when the function representing the primary tooth surface is in parametric form. However, the zone of possible conjugate action can always be found explicitly point by point.

Given a form-cutting method by which we wish to produce the teeth of the primary gear member, there will always be some arbitrary parameters. In spiroid worms, which are cut with a taper-helical motion of constant lead, we can vary the lead, taper angle, position, and the pressure angles of the teeth. The tooth action law helps us choose these parameters in a way which will insure conjugate action.

Referring now to the specific right-angle skew-axis case, Figure 3, let us say that conjugate action is desired at some arbitrary point (called the "pitch point") which is designated by giving numerical values to its x, y, z coordinates. If the center distance C and speed ratio K are given, then the relative helices are determined, and one of them passes through the pitch point p. Its direction in space is the same as that of the relative velocity vector given by equation (1).

It is convenient to relate the relative velocity vector to a new Cartesian coordinate system whose origin is at the arbitrary point p, as shown in Figure 7.

Point p is the arbitrary point, and r is the radial distance from axis #1 to this point. The new coordinate system is designated by three mutually perpendicular vectors $\vec{i}$, $\vec{a}$, and $\vec{b}$, originating from p and defined as follows:

$\vec{i}$ is the x-unit vector, and is therefore parallel to axis #1.

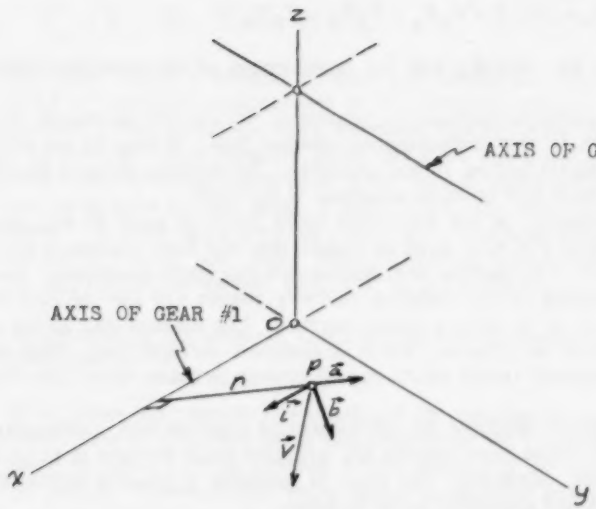


Figure 7

Transformation of Coordinates

$\vec{a}$ is a unit vector in the direction of the radial line through p, and is therefore perpendicular to axis #1.

$\vec{b}$ is a unit vector at right angles to both $\vec{i}$ and $\vec{a}$, defined by $\vec{b} = \vec{a} \times \vec{i}$, and is therefore at right angles to both axis #1 and the radial line r.

These three new unit coordinate vectors can be expressed in terms of the x, y, z coordinates of the point p as follows:

$$\vec{i} = \vec{i}$$

$$\vec{a} = \frac{1}{r}(y\vec{j} + z\vec{k})$$

$$\vec{b} = \frac{1}{r}(z\vec{j} - y\vec{k})$$

Let V_x, V_a, and V_b be the components of the relative velocity vector in these coordinate directions. Then

$$V_x = \vec{V} \cdot \vec{i} = C - z$$

$$V_a = \vec{V} \cdot \vec{a} = \frac{xz}{r}$$

$$V_b = \vec{V} \cdot \vec{b} = Kr - \frac{xy}{r}$$

The relative velocity vector at p is tangent to the surface of revolution (or pitch surface) generated by the relative helix through p. Let τ be the taper angle of this surface at the point p, as defined in Figure 8. This instantaneous taper angle of the pitch surface at p is given by

$$(8) \qquad \tan\tau = \frac{V_a}{V_x} = \frac{xz}{r(C-z)}$$

The relative velocity vector is tangent to a cone of this taper angle, whose axis is on the x-axis, and which contains p. Hence it is co-directional with some helix of constant lead drawn in the cone. This helix must have a lead L which is related to the axial and tangential components of the relative velocity vector as follows,

$$(9) \qquad L = (2\pi r)\frac{V_x}{V_b} = \frac{2\pi(C-z)}{K - \frac{xy}{x^2}}$$

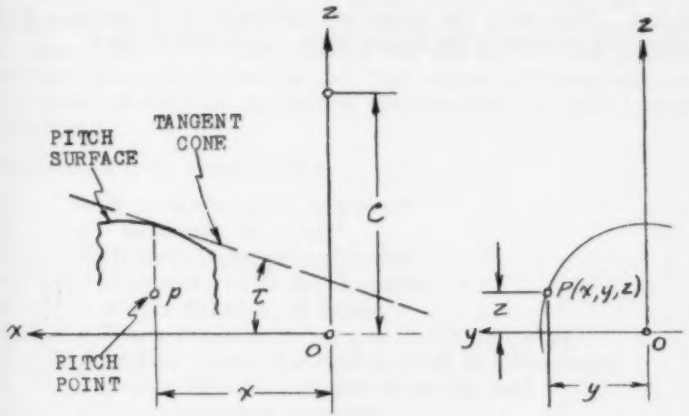


Figure 8

Pitch Surface and Tangent Cone

Equations (8) and (9) are fundamental formulas in spiroid gearing. A spiroid worm which has a taper-helical element of this taper angle and lead in a conical reference surface containing p will have possible conjugate action at p regardless of other tooth form parameters. In particular, if both sides of the worm thread are generated along this tapered helix, they will have conjugate action at p. The amount of conjugate action on the remainder of the thread surfaces depends on the pressure angles and on the location of the pitch p. In general it is better to have the pitch point on the outside of the worm blank as this minimizes the amount of inactive fillet generated in the mating gear.

It is interesting to note that conventional cylindrical worms are but a special case of the gearing defined by these formulas. Here the pitch point is on the common perpendicular between the axes of the gears (the z-axis). Thus $x = 0$, $y = 0$. Substituting in equations (8) and (9) gives $\tau = 0$ and $L = 2\pi(C-z)/K$. But this means the pitch surface is cylindrical and the lead is exactly that of a conventional worm whose pitch radius is z. Furthermore, if $x = 0$, $y \neq 0$, and $z = \text{constant}$, we get the same taper angle and lead from the formulas. This means that pitch points chosen on a line in the y-z plane, parallel to the y-axis, all give the same worm thread parameters. Therefore the zones of contact for both sides of the worm threads intersect along this line regardless of pressure angles. But this is simply the well-known pitch line of conventional worm gearing.

Spiroid gears should be thought of as a generalized form of worm gearing. They apply the foregoing principles to the purpose of extending and improving the nature of the tooth contact lines.

10

Design of Orifices for Hydraulic Shock Absorbers

By T. E. Lohr
General Motors Corporation
Detroit

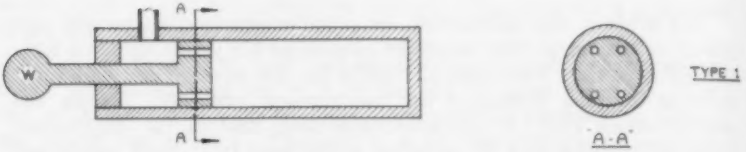
The modern day automobile, as it increases in size, weight, and power, is becoming what would be classified ten years ago as a full-fledged race car. This type of vehicle for our present day highways imposes serious problems for the chassis engineer. Among the most pressing are: (1) high speed side sway control, (2) longitudinal "dipping" control, and (3) isolation of the body from road bumps and noises. A great deal of development has been done in suspension systems such as the "air-ride," the full and partial torsion ride, and the torsion air ride, to attempt to improve roadability. However, it is apparent that regardless of which type of suspension system is used, the success or failure of that system depends either partly or wholly on the shock absorber selected. In the infancy of the automobile, shock absorbers were designed by trial and error until a suitable unit was chosen for the vehicle. Today, this field has become an exact science.

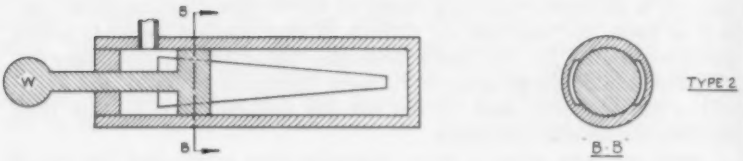
Until recently, when a large manufacturer announced the use of a new aircraft type shock absorber known as the "oriflow," we were still applying old concepts of design. This new method of ride control opened the eyes of the industry and has led to even greater improvements in shock absorption. This article will deal with two basic types of orifices for future consideration in ride control developments.

Symbols used in the computations:

a	=	Total orifice area, in inches^2
A	=	Piston area, in inches^2
C	=	Orifice discharge coefficient
d_c	=	Diameter of cylinder, in inches
D	=	Piston diameter, in inches
D_o	=	Piston diameter at start of travel, in inches
E	=	Kinetic energy absorbed by fluid, in foot-pounds
f	=	Acceleration, in foot-pounds per second^2
F	=	Force on piston, in pounds
g	=	Acceleration of gravity = 32.2 feet per second^2
P	=	Pressure, in pounds per inch^2

S = Piston travel from start, in inches
S_{max} = Full piston stroke, in inches
v = Velocity of piston, in feet per second
V = Velocity of fluid, in feet per second
W = Weight of body to be decelerated, in pounds
γ = Density of fluid in cylinder, in pound per foot^3 (hydraulic fluid = 69.4 lbs/ft^3)
M = Mass = W/g





Actually, there are five types of orifices although only two are in widespread use. It is these two that we are concerned with.

EQUATIONS:

(1) $$W\frac{dv}{v} = \gamma C^2 A^3 \, ds/3456\, a^2$$

(2) $$P = \frac{\gamma C^2 A^2 V^2}{288\, g a^2}$$

(3) Type 1 $$\text{Log}_e \frac{v_0}{v} = \frac{\gamma C^2 A^2 S}{3456\, a^2 W}$$

(4) Type 2 $$W\text{Log}_e \frac{v_0}{v} = \gamma \frac{C^2 A^3}{13824\, d_c\, K(t_{n\ 2})^2} \left[- \frac{\cos(\phi/2)}{(\phi/2)} - \frac{\phi}{2} + \frac{(\phi/2)^3}{3L3} - \frac{(\phi/2)^5}{5L5} + \cdots \right]_{\phi_0}^{\phi}$$

(5) $$P = F/A = Wf/gA$$

To illustrate the application of the equations, a typical shock absorber problem will be considered. The various orifice sizes will be calculated and the performance will be analyzed in terms of cylinder pressure and piston velocity versus piston stroke.

Problem:

A 1000 pound weight traveling at a rate of 100 feet per second is to be brought to rest in a distance of three feet. Space limitations require that the piston diameter be approximately six inches.

TYPE 1 ORIFICE

Assume:

$a = 3$ inches^2 — Substituting in equation 3, we have:

$$\text{Log}_e(100/v) = 76 \times 1.2^2 \times (0.7854 \times 6^2)^3 s/(3456 \times 3^2 \times 1000),$$

from which it is possible to solve for values of v for corresponding values of s.

By using equation 2, also, the following tape has been prepared:

S	0	2	4.5	9	18	27	36
v	100	85.2	70	49	24	11.7	5.75
P	10500	8260	5600	2520	604	153	34.7

TYPE 2 ORIFICE

Assume total orifice areas to vary between 8 inches^2 and 0.25 inches^2, divided among three orifices. The piston diameter is six inches, and the depth of slot is 0.75 inches.

The central angle subtended by each orifice will then vary between 60.4 and 1.9 degrees, for which the corresponding chords will be 3.4 inches and 0.111 inches. This will give K a value of 0.0915. By use of equations (4) and (2), the following table is computed: Note that in equation (4):

T = Depth of slot = 0.75 inch
N = Number of slots = 3
ϕ = Central angle subtended by slot radians

$$K = \frac{X_0 - X}{S}$$ where:

X_0 = Maximum width of slot
X = Width at given station
S = Piston stroke from initial position.

S	0	3	9	12	18	22.5	27	30	36
X	3.4	3.13	2.59	2.30	1.75	1.34	0.93	0.66	0.11
ϕ	60.4	55.2	44.4	40	30	25.3	15.82	11.2	1.89
a	8.00	7.31	5.88	5.30	3.97	3.03	2.09	1.49	0.25
v	100	93.2	85.7	78	64.6	50.7	24.9	12.1	0
P	1477	1530	2000	2040	2570	2713	1365	645	0

EFFECT OF FRICTION AND SPRING FORCES

In the foregoing discussion, piston velocity is assumed to be affected only by orifice size. This is the ideal case assuming that friction is negligible; however, any or all of the following forces may exist: (1) mechanical friction, (2) restoring forces exerted by springs for returning the piston to zero position, (3) packing friction, (4) viscous friction, (5) mechanism weights when arranged in a vertical direction, and (6) hydraulic fluid elasticity. Viscous friction and fluid elasticity will be taken up later. The method of including constant and variable resistance will be considered here:

Kinetic energy to be dissipated = $Wv^2/2g$.
Frictional energy (mechanical and packing) = $K_1\, s/12$
where K_1 = frictional load.
Spring energy = $\int (S_0+S)\, K_2\, d_s /12$ where
S_0 = pre-compression, in inches
K_2 = spring constant, in inch/pounds.

For a constant decleration orifice, merely subtract the frictional energy from the kinetic energy and solve for the necessary orifice area "a". For all other cases, the procedure is as follows:
The energy equation is

$$W(v_0^2 - v^2)\, 2g = \int_0^S \frac{\gamma C^2 v^2 A^3 d_s}{3456\, g a^2} + \int_0^S K_1 \frac{d_s}{12} + \int_0^S \frac{(S_0+S)K_2 d_s}{12}$$

Differentiate the energy equation above:

$-W_v d_v/g = K_3 v^2 A^3\, d_s/a^2 + (K_1+K_2 S_0)\, d_s/12 + K_2 S d_s/12$ where

$K_3 = \gamma C_2/3456\, g$.

Divide thru by d_s:

$$-\frac{W_v}{g}\frac{d_v}{d_s} = K_3 v^2 A^3/a^2 + (K_4 + K_2 S)/12 \text{ where } K_4 = K_1 + K_2 \text{ so let}$$

$$v = \sqrt{y} \text{ then } \frac{W}{2g} \cdot \frac{d_y}{d_s} \cdot \frac{-K_3 A_3 y}{a^2} = \frac{K_4+K_2 S}{12} .$$

This equation is a linear, first order differential equation, the solution of which is

$$y = v^2 = e^{-\int K_3 2gA^3 d_s/(Wa^2)} \left\{ -\int e^{\int K_3 2gA^3 d_s/Wa^2} 2g(K_4 + K_2 S) d_s/12\, w + c \right\}$$

from which it can be seen that for $S = 0$, $c = v^2 0$. It may be noted that the value of $\int K_3 2gA_3/(W_a{}^2) d_s$ may be computed directly by multiplying the right side of (3), (4), and (5) by 2 W.

In order to obtain an idea of the comparative effect of frictional forces, their effect on piston speed will be considered.

Assume that there is a constant mechanical friction force of 250 pounds due to a coefficient of friction of 0.25, also, that there is a constant packing force of 960 pounds. This value is determined by multiplying the highest pressure to be encountered by 0.2. There is a spring retarding force which is from 1000 pounds at the beginning of the stroke to 2000 pounds at the end of the stroke. Let the problem be to determine the velocity of the piston at 18 inches or 50 per cent of the stroke. By equation (6):

$$(6) \qquad \int_0^S K_3 \frac{2gA^3}{W_a{}^2} d_s = \frac{2w}{w} \operatorname{Log}_e (v/v_0) = -0.391. \quad \text{Also}$$

$$K_4 = K_1 + K_2 S_0 = 250 + 960 + 1000 = 2210,$$

substituting in equation (6) and solving step by step integration, v = 81 feet/second compared to 82.2 feet/second calculated without including frictional effects. In this case, the effect of friction is negligible, but it cannot be assumed that the difference will be negligible in every case. Usually the friction effect may be neglected in making the first approximation for orifice size.

EFFECT OF VISCOSITY

Frictional losses due to fluid viscosity are small in shock-absorber design. In the example chosen for this article, assume that the piston has traveled 2 inches. The coefficient of viscosity, $\mu = 0.00003$ pounds per second/foot.[2] Assume an orifice of type 2 is used:

$$F_v = \mu a\, v/y = \mu a\, A_v/(ay) \quad \text{where a = circumferential area of the piston}$$

then $$F_v = \frac{3 + 10^{-5} \times 6\pi \times 2 \times 0.7854 \times 6^2}{144 \times 0.75/12} \cdot \frac{v}{a}$$

F_v, for a velocity of 100 feet/sec will then be approximately 0.5 pounds, which is negligible compared to a hydraulic force averaging about 40,000 pounds.

EFFECT OF FLUID ELASTICITY

The average volume modulus of elasticity of water at an operating pressure of about 4000 p.s.i.is 330,000 p.s.i. Also E_v = volume modulus of elasticity in p.s.i.; P = Pressure, in P.S.I.; V = Volume, in ft.^{3}

If maximum pressure developed using type 1 orifice is 10,500 P.S.I. and the volume of the fluid is 0.59 feet^3, then V = 0.59 x 10,500/330,000 = 0.0188 cubic foot compression. This amounts to a linear displacement (using 0.5 foot diameter) of 0.0957 feet. The work absorbed by this compression is 0.5 x 0.0957 x 10500 = 503 foot/pounds, which amounts to 3 per cent of the original kinetic energy. However, when it is considered that the pressure is reduced to zero by the time that the piston reaches the end of its travel, it may be seen that the energy of elasticity has been reduced considerably.

Continuing:

Bernoulli's Theorem: - $$\frac{144p}{\gamma} = \frac{C^2V^2}{2g} \quad \text{or} \quad P = \frac{C_2V^2\gamma}{288\ g}.$$

hence the force on the piston = F = $A_p = \frac{\gamma C^2 V^2 A}{288\ g}$ while the work done by the piston during its stroke = E = $\int_0^s Fd_s/12$.

This equals the change in kinetic energy of the mass to be stopped. Hence:

$$\frac{W(v_o^2 - v^2)}{2g} = \int_0^s \frac{\gamma C^2 V^2 A d_s}{3456\ g}.$$

By continuing equation - $V = A_v/a$, therefore, where v_o = velocity of weight (W) at start of stroke

$$\frac{W(v_o^2 - v^2)}{2g} = \int_0^s \frac{\gamma C^2 V^2 A^3 d_s}{3456\ ga^2}.$$

Differentiating and dividing three by v^2

$$-W\frac{dv}{v} = \frac{\gamma C^2 A^3 d_s}{3456 a^2},$$

also, we have from above $P = \gamma C^2 A^2 v^2 / 288\, g a^2$.

TYPE 2 ORIFICE

The usual method of solving for velocity is to divide the stroke into a sufficient number of small segments and compute the pressure over the first segment, using $P = \frac{2c^2 A^2 v^2}{288\, g a^2}$ and assuming that $v = v_o$ is constant over the whole segment. The deceleration then is computed using $P = wf/gA$. The velocity for the next segment is found, remembering that $v^2 = v_o^2 - 2fs/12$. This process is repeated until the velocity is found for every segment. The disadvantages of this method are as follows: (1) The fewer the segments taken, the more the point of maximum pressure is moved toward the beginning of the stroke. By integration, the actual point of maximum pressure is found to be nearer the end of the stroke. (2) Any error at the beginning of the calculations will necessitate starting over at the point of error; moreover, the error is cumulative. (3) And many times, after all of the work of computation has been carried out, it will be found that the entire stroke has not been used efficiently, in that v becomes negligible too soon along the stroke (thus introducing higher pressures than are necessary) or that v is too high at the end of travel, causing shock absorption to be inadequate.

If the width of the slot is described by the following expressions: $(X_o - x)/S = K$ where $K/2$ also $\text{Tan}\ (\phi/2)$ and $\sin(\phi/2) = xd_c = (X_o - K_s)/d_c$ and $K_s = X_o - d_c \sin(\phi/2)$ and $ds = -d_c \cos(\phi/2)d\phi/(2K)$. Also, let

n = number of slots
t = thickness of slot, in inches
ϕ = angles of slot, radius

Then: $a = d_c \phi t\, n/2$.

Substituting: $$\frac{wdv}{v} = \frac{2c^2 A^3 d_c \cos(\phi/2)d\phi}{6912K (d_c/2)^2 t^2 n^2 \phi^2}$$

Integrating we have:

$$w \log_e \frac{v}{v_o} = \frac{\gamma c^2 A^2}{13824\, d_c K (t_n/2)^2} \cdot \left[\frac{-\cos(\phi/2)}{(\phi 2)} - \frac{\phi}{2} + \frac{(\phi/2)^3}{3L3} - \frac{(\phi/2)^5}{5L5} + - + - -\right]_{\phi_o}^{\phi}.$$

The remainder of the calculations would then be similar to those described earlier in the article.

INDUSTRIAL MATHEMATICS

Industrial Mathematics is published annually by Wayne State University Press for the Industrial Mathematics Society. The advance subscription price is $3.00 per annual volume. Back issues are available at $3.00 per copy. Orders accompanied by check or money order should be addressed to Wayne State University Press, 4841 Cass Avenue, Detroit 2, Michigan.

Industrial Mathematics contains original articles dealing with the application of mathematics in industry. Persons having such manuscripts for publication should submit them in duplicate to the Editor, *Industrial Mathematics*, 100 Farnsworth Street, Detroit 2, Michigan. Any questions on editorial matters should be similarly addressed.

Manuscripts should be typed double-spaced on one side of $8\frac{1}{2}$ by 11 inch paper in complete and finished form precisely as they are to appear, and sent—packed flat—to the Editor. Where possible, illustrations should be drawn in India ink as boldly as possible. Figures should be numbered consecutively in arabic numerals. Tables should be numbered in arabic numerals, and both figures and tables should have captions. Ten free reprints of each article will be furnished authors.

WAYNE STATE UNIVERSITY PRESS; Publisher
4841 Cass Avenue
Detroit 2, Michigan
U.S.A.